

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 08-249347
 (43)Date of publication of application : 27.09.1996

(51)Int.Cl. G06F 17/30

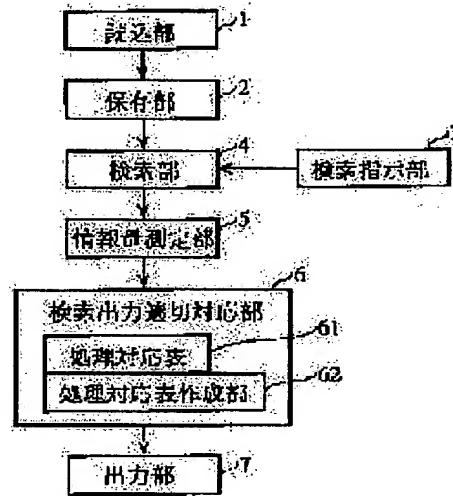
(21)Application number : 07-051345 (71)Applicant : MATSUSHITA ELECTRIC IND CO LTD
 (22)Date of filing : 10.03.1995 (72)Inventor : TERAMURA KANAKO KURAMITSU ASAMI

(54) METHOD AND DEVICE FOR DATA RETRIEVAL

(57)Abstract:

PURPOSE: To easily confirm the retrieval result by making a different display from display of a measured value less than a specific value when a measured value obtained in an information amount measurement step exceeds the specific value.

CONSTITUTION: A retrieval part 4 detects data meeting retrieval conditions, one by one, in a storage part 2 and sends their titles to an information amount measurement part 5. When all data meeting the retrieval conditions are detected, the information amount measurement part 5 finds the number of the titles of found candidate data as retrieval results. And, it is judged whether or not rereretrieval is needed from the found value by referring to a process correspondence table. When the rereretrieval such as input is necessary, that is indicated to a retriever, but when not, an output format of the data as the retrieval result is so selected as to match with the process correspondence table. Consequently, when the number of the titles of the candidate data is 10-50, the titles of all the candidate data reported to the information amount measurement part 5 are outputted and when <10, data corresponding to the titles are all read out of the storage part 2 and outputted.



LEGAL STATUS

[Date of request for examination] 30.10.2000

[Date of sending the examiner's decision of rejection] 16.12.2003

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平8-249347

(43)公開日 平成8年(1996)9月27日

(51)Int.Cl. ⁶ G 0 6 F 17/30	識別記号 9194-5L 9194-5L 9194-5L 9194-5L	序内整理番号 F I G 0 6 F 15/403 15/40 15/403 3 4 0 B	技術表示箇所 3 7 0 A 3 7 0 G 3 4 0 Z 3 4 0 B
---	--	---	--

審査請求 未請求 請求項の数26 OL (全27頁)

(21)出願番号 特願平7-51345	(71)出願人 000005821 松下電器産業株式会社 大阪府門真市大字門真1006番地
(22)出願日 平成7年(1995)3月10日	(72)発明者 寺村 佳奈子 大阪府門真市大字門真1006番地 松下電器 産業株式会社内
	(72)発明者 倉光 麻美 広島市東区光町1丁目12-20 ヒロソー栄 泉ビル3階 株式会社松下電器情報システ ム広島研究所内
	(74)代理人 弁理士 中島 司朗

(54)【発明の名称】 データ検索方法及びデータ検索装置

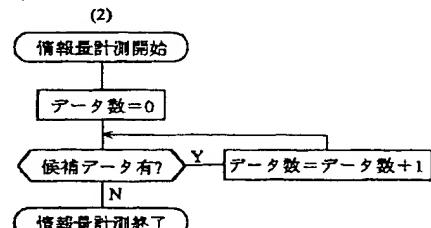
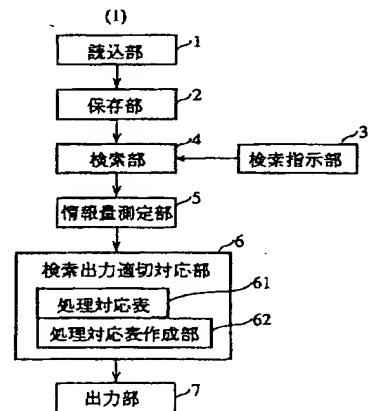
(57)【要約】

【目的】 電子化された上記憶保管されている大量のデータや文字放送等での検索効率の改善を図る。

【構成】 (1) 一応検索されたデータの情報量をもとに、再検索の必要性やその手法を検索者に提供したりあるいは検索結果の出力方法等を変更する。

(2) データの利用頻度等があらかじめ判明している場合には、検索やその後の出力にそれを反映させる。

(3) 検索者の記憶する定型文書中の単語、文字等の読み取りを検索に利用する。



検索者の指示により、保存、放送、提供等されているデータから候補データを読み出す検索手段と、

前記検索手段において読み出された全データの情報量を計測する情報量測定手段と、

前記情報量測定手段の測定値があらかじめ決められた制限値を超えているか否かを判定し、制限値を超えている場合には、前記検索手段により読み出されたデータを所定の規則により分類する検索データ分類手段と、

前記検索データ分類手段にて分類された各分類毎に、出力を行うか否かを所定の手順で判断する出力データ分類選択手段とを有していることを特徴とするデータ検索装置。

【請求項1】 文字、表、グラフ、図形等で構成されるデータを検索し、結果をデータファイル名、データ内容、データファイル属性、データ特徴などの形式で出力するデータ検索方法であって、
検索者の指示により、保存、放送、提供等されているデータから候補データを読み出す検索ステップと、
前記検索ステップにおいて読み出された全データの情報量を測定する情報量測定ステップと、
前記情報量測定ステップによる測定値が所定の値を越えた場合には、所定値未満の場合とは別形式の出力を行う、検索者に出力方法の選択指示をあおぐ、検索者に他の検索内容による再検索を要請するの少なくも一を行う検索結果出力適切対応ステップとを有していることを特徴とするデータ検索方法。

【請求項2】 文字、表、グラフ、図形等で構成されるデータを検索し、結果をデータファイル名、データ内容、データファイル属性、データ特徴などの形式で出力するデータ検索装置であって、
検索者の指示により、保存、放送、提供等されているデータから候補データを読み出す検索手段と、
前記検索手段において読み出された全データの情報量を測定する情報量測定手段と、
前記情報量測定手段による測定結果が所定値を越えた場合には、所定値未満の場合とは別形式の出力を行う、検索者に出力方法の選択指示をあおぐ、検索者に他の検索内容による再検索を要請するの少なくも一を行う検索結果出力適切対応手段とを有していることを特徴とするデータ検索装置。

【請求項3】 文字、表、グラフ、図形等で構成されるデータを検索し、結果をデータファイル名、データ内容、データファイル属性、データ特徴などの形式で出力するデータ検索方法であって、
検索者の指示により、保存、放送、提供等されているデータから候補データを読み出す検索ステップと、
前記検索ステップにおいて読み出された全データの情報量を測定する情報量測定ステップと、
前記情報量測定ステップの測定値があらかじめ決められた制限値を超えているか否かを判定し、制限値を超えている場合には、前記検索ステップにより読み出されたデータを所定の規則により分類する検索データ分類ステップと、
前記検索データ分類ステップにて分類された各分類毎に、出力を行うか否かを所定の手順で判断する出力データ分類選択ステップとを有していることを特徴とするデータ検索方法。

【請求項4】 文字、表、グラフ、図形等で構成されるデータを検索し、結果をデータファイル名、データ内容、データファイル属性、データ特徴などの形式で出力するデータ検索装置であって、

【請求項5】 文字、表、グラフ、図形等で構成されるデータを検索し、結果をデータファイル名、データ内容、データファイル属性、データ特徴などの形式で出力するデータ検索方法であって、
検索者の指示により、保存、放送、提供等されているデータから候補データを読み出す検索ステップと、
前記検索ステップにおいて読み出された全データの情報量を測定する情報量測定ステップと、

前記情報量測定ステップの測定値があらかじめ決められた制限値を超えているか否かを判定して、制限値を超えている場合には、前記検索ステップにより読み出されたデータについて所定のデータファイル属性あるいはデータ特徴等により分類する検索データ分類ステップと、
前記検索データ分類ステップにて分類された各分類毎に、同じ分類のデータが複数存在することとなる分類のデータのみ出力するよう制御する同一分類データ出力制御ステップとを有していることを特徴とするデータ検索方法。

【請求項6】 文字、表、グラフ、図形等で構成されるデータを検索し、結果をデータファイル名、データ内容、データファイル属性、データ特徴などの形式で出力するデータ検索装置であって、
検索者の指示により、保存、放送、提供等されているデータから候補データを読み出す検索手段と、
前記検索手段において読み出された全データの情報量を測定する情報量測定手段と、
前記情報量測定手段の測定値があらかじめ決められた制限値を超えているか否かを判定し、制限値を超えている場合には、前記検索手段により読み出されたデータについて特定のデータファイル属性あるいはデータ特徴等により分類する検索データ分類手段と、

前記検索データ分類手段にて分類された各分類毎に、同じ分類のデータが複数存在することとなる分類のデータのみ出力するよう制御する同一分類データ出力制御手段とを有していることを特徴とするデータ検索装置。

【請求項7】 文字、表、グラフ、図形等で構成されるデータを検索し、結果をデータファイル名、データ内

容、データファイル属性、データ特徴などの形式で出力

するデータ検索方法であって、
検索者の指示により、保存、放送、提供等されているデータから候補データを読み出す検索ステップと、
前記検索ステップにて読み出された全データの情報量を測定する情報量測定ステップと、
前記情報量測定ステップの測定値があらかじめ決められた制限値を超えているか否かを判定し、制限値を超えている場合には、前記検索ステップにより読み出されたデータについて所定のデータファイル属性あるいはデータ特徴等により分類する分類ステップと、
前記分類ステップの結果、同じ分類のデータが存在しないデータのみ出力するよう制御する非同一分類データ出力制御ステップとを有していることを特徴とするデータ検索方法。

【請求項 8】 文字、表、グラフ、図形等で構成されるデータを検索し、結果をデータファイル名、データ内容、データファイル属性、データ特徴などの形式で出力するデータ検索装置であって、
検索者の指示により、保存、放送、提供等されているデータから候補データを読み出す検索手段と、

前記検索手段において読み出された全データの情報量を測定する情報量測定手段と、
前記情報量測定手段の測定値があらかじめ決められた制限値を超えているか否かを判定し、制限値を超えている場合には、前記検索手段により読み出されたデータについて特定のデータファイル属性あるいはデータ特徴等により分類する検索データ分類手段と、
前記検索データ分類手段にて分類された各分類毎に、同じ分類のデータが存在しないデータのみ出力するよう制御する非同一分類データ出力制御手段とを有していることを特徴とするデータ検索装置。

【請求項 9】 文字、表、グラフ、図形等で構成されるデータを検索し、結果をデータファイル名、データ内容、データファイル属性、データ特徴などの形式で出力するデータ検索方法であって、
検索者の指示により、保存、放送、提供等されているデータから候補データを読み出す検索ステップと、

前記検索ステップにおいて読み出された全データの情報量を測定する情報量測定ステップと、
前記情報量測定ステップにより測定された情報量をあらかじめ定めたいずれの範囲内であるかを判定する範囲判定ステップと、

前記範囲判定ステップの判定した範囲に応じて前記検索ステップから読み出されたデータの出力方法若しくは出力内容の少くとも一を制御する出力制御ステップとを有していることを特徴とするデータ検索方法。

【請求項 10】 文字、表、グラフ、図形等で構成されるデータを検索し、結果をデータファイル名、データ内容、データファイル属性、データ特徴などの形式で出力するデータ検索装置であって、

検索者の指示により、保存、放送、提供等されているデータから候補データを読み出す検索手段と、
前記検索手段において読み出された全データの情報量を測定する情報量測定手段と、
前記情報量測定手段により測定された情報量の値があらかじめ決められたどの範囲であるかを判定する範囲判定手段と、
前記範囲判定手段の判定した範囲に応じて前記検索手段から読み出されたデータの出力方法若しくは出力内容の少くとも一を制御する出力制御手段とを有していることを特徴とするデータ検索装置。

【請求項 11】 文字、表、グラフ、図形等で構成されるデータを検索し、結果をデータファイル名、データ内容、データファイル属性、データ特徴などの形式で出力するデータ検索方法であって、
検索者の指示により、保存、放送、提供等されているデータから候補データを読み出す検索ステップと、

前記検索ステップにおいて読み出された全データの情報量を測定する情報量測定ステップと、
前記情報量測定ステップの測定値があらかじめ設定された制限値を超えているか否かを判定し、制限値を超えている場合には、前記検索ステップにより読み出されたデータをあらかじめ決められた規則に従って並び替えて出力するよう制御する出力順制御ステップとを有していることを特徴とするデータ検索方法。

【請求項 12】 文字、表、グラフ、図形等で構成されるデータを検索し、結果をデータファイル名、データ内容、データファイル属性、データ特徴などの形式で出力するデータ検索装置であって、
検索者の指示により、保存、放送、提供等されているデータから候補データを読み出す検索手段と、
前記検索手段において読み出された全データの情報量を計測し、その値があらかじめ設定された制限値を超えているか否かを判定する情報量測定手段と、
前記情報量測定手段により測定された情報量が制限値を超えている場合、前記検索手段により読み出されたデータをあらかじめ決められた条件で並び替え、その順で出力するよう出力手段に指示する出力順制御手段とを有していることを特徴とするデータ検索装置。

【請求項 13】 文字、表、グラフ、図形等で構成されるデータを検索し、結果をデータファイル名、データ内容、データファイル属性、データ特徴などの形式で出力するデータ検索方法であって、
検索者の指示により、保存、放送、提供等されているデータから候補データを読み出す検索ステップと、
前記検索ステップにおいて読み出された各データの情報量を順に確認しつつ計測する情報量順計測ステップと、
前記情報量順計測ステップの結果に基づきそれまでの総和があらかじめ定められた制限値を超えているか否かを判定する情報量測定ステップと、

前記情報量測定ステップにより測定された情報量の総和が制限値を超えた場合には、制限値に該当するデータ以降に読み出されたデータは出力しない若しくは制限値を超えた時点で検索を中止するよう制御する出力検索制限制御ステップとを有していることを特徴とするデータ検索方法。

【請求項14】 文字、表、グラフ、図形等で構成されるデータを検索し、結果をデータファイル名、データ内容、データファイル属性、データ特徴などの形式で出力するデータ検索装置であって、

使用者の操作に基づく検索起動指示により、保存、放送、提供等されているデータから候補データを読み出す検索手段と、

前記検索手段において読み出された各データの情報量を順に測定する情報量測定手段と、

前記情報量測定手段の測定した総和があらかじめ設定された制限値を超えているか否かを判定し、制限値を超えた場合には、それ以降に前記検索手段により読み出されたデータは出力しない若しくは制限値を超えた時点で検索を中止するよう制御する出力検索制限手段とを有していることを特徴とするデータ検索装置。

【請求項15】 前記範囲判定ステップにて所定の範囲にあると判断されたならば、前記検索ステップにて検索された候補データの中に画像データが含まれているか否かを判断する画像データ存在判断ステップを有し、

前記範囲判定ステップは更に、同じく画像データが存在すると判断されたならば、そのデータ量を測定する画像データ量測定小ステップを有し、

前記出力制御ステップは更に、

前記画像データ存在判断ステップにて存在すると判断されたならば、画像データを先に出力されるよう候補データの配列を調整する画像データ優先配列小ステップと、前記画像データ量測定小ステップにて画像データ量が所定値以上と判断されたならば、画像データが出力される際にこれを縮尺して、同一表示面に複数の画像データが同時に出力されるよう制御する画像データ出力表示制御小ステップとを有していることを特徴とする請求項9記載のデータ検索方法。

【請求項16】 前記範囲判定手段にて所定の範囲にあると判断されたならば、前記検索手段にて検索された候補データの中に画像データが含まれているか否かを判断する画像データ存在判断手段を有し、前記範囲判定手段は更に、同じく画像データが存在すると判断されたならば、そのデータ量を測定する画像データ量測定部を有し、

前記出力制御手段は更に、

前記画像データ存在判断手段にて存在すると判断されたならば、画像データを先に出力されるよう候補データの配列を調整する画像データ優先配列部と、

前記画像データ量測定部にて画像データ量が所定値以上

と判断されたならば、画像データが出力される際にこれを縮尺して、同一表示面に複数の画像データが同時にに出力されるよう制御する画像データ出力表示制御部とを有していることを特徴とする請求項10記載のデータ検索装置。

【請求項17】 検索装置による検索の対象とされている、そして保存されているデータ若しくは保存するデータに経験則、データの意味内容、キーワード等から定められた本来の検索の目的とされる確率についての情報を検索に先立って付加する利用度情報付加ステップを有し、

前記出力制御ステップは、前記範囲判定ステップの判定した範囲に応じて前記検索ステップから読み出されたデータの出力方法若しくは出力内容を制御するに際して、前記利用度情報付加ステップにて付加された情報を反映させる付加利用度情報反映输出力制御小ステップを有していることを特徴とする請求項9記載のデータ検索方法。

【請求項18】 検索装置による検索の対象とされている、そして保存されているデータ若しくは保存するデータに経験則、データの意味内容、キーワード等から定められた本来の検索の目的とされる確率についての情報を検索に先立って付加する利用度情報付加手段を有し、

前記出力制御手段は、前記範囲判定手段の判定した範囲に応じて前記検索手段から読み出されたデータの出力方法若しくは出力内容を制御するに際して、前記利用度情報付加手段にて付加された情報を反映させる付加利用度情報反映输出力制御部を有していることを特徴とする請求項10記載のデータ検索装置。

【請求項19】 前記出力検索制限御ステップにより出力を制限されないこととなったデータについては、あらかじめ定められた所定の規則に並び替えて出力するよう制御する出力順制御ステップを有していることを特徴とする請求項14記載のデータ検索方法。

【請求項20】 前記出力検索制限手段により出力を制限されないこととなったデータについて、あらかじめ定められた所定の規則に並び替え、その順で出力するよう出力手段を制御する出力順制御手段を有していることを特徴とする請求項15記載のデータ検索装置。

【請求項21】 記事の意味内容から定まる記事の欄の分類と該記事の記載されている位置若しくはこれに併せての記載部分の形状、寸法、記載部やその近辺の所定の記号、図形、写真、枠等記載部分を特定するのに役立つ視覚的特徴との間に一定の関係の存在する定型文書から目的とする記載を検索する方法において、

検索目的たる記事等が記載されている少くも一の種類の定型文書について、その定型性についての情報を検索に先立ってあらかじめ入力しておく定型性情報入力ステップと、

検索目的たる記事の意味内容から定まる記事の欄の分類若しくはこれに併せての該記事の記載されている定型文

書の種類を検索キーとして入力する欄定型情報入力ステップと、

検索目的たる記事中の文章、単語、文字の少くも一方について、少くも一個を検索キーとして入力する若しくはこれに併せて記載位置、近辺の所定の記号等をも検索キーとして入力する記事情報入力ステップとを有していることを特徴とする定型文書用記事中の文書検索方法。

【請求項22】 記事の意味内容から定まる記事の欄の分類と該記事の記載されている位置若しくはこれに併せての記載部分の形状、寸法、記載部やその近辺の所定の記号、図形、写真、枠等記載部分を特定するのに役立つ視覚的特徴との間に一定の関係の存在する定型文書から目的とする記載を検索する装置において、

検索目的たる記事等が記載されている少くも一の種類の定型文書について、その定型性についての情報を検索に先立ってあらかじめ入力され記憶している定型性情報保持手段と、

検索目的たる記事の意味内容から定まる記事の欄の分類若しくはこれに併せての該記事の記載されている定型文書の種類を検索キーとして入力する欄定型情報入力手段と、

検索目的たる記事中の文章、単語、文字の少くも一について、少くも一個を検索キーとして入力する若しくはこれに併せて記載位置、近辺の所定の記号等をも検索キーとして入力する記事情報入力手段とを有していることを特徴とする定型文書用記事中の文書検索装置。

【請求項23】 相前後する単語間に余白を設ける言語で記載された文を対象として、文中に存在する単語を検索キーとして、所望するデータの検索を行なうデータ一検索方法であって、

検索対象の文中の単語間の余白から文中の各単語の長さを求める単語長検出手段と、

検索キーの単語長を所定の手順で求める検索キー単語長検出手段と、

前記両検出手段の結果をもとに、検索対象の文から検索キーの単語と同じ長さの単語を見出す候補単語検出手段と、

前記候補単語検出手段により見出された候補単語に対して、文字認識を行なって検索キーの単語か否かの判断を行なう最終文字認識手段とを有していることを特徴とするデータ検索方法。

【請求項24】 相前後する単語間に余白を設ける言語で記載された文を対象として、文中に存在する単語を検索キーとして、所望するデータの検索を行なうデータ一検索装置であって、

検索対象の文中の単語間の余白から文中の各単語の長さを求める単語長検出手段と、

検索キーの単語長を所定の手順で求める検索キー単語長検出手段と、

前記両検出手段の結果をもとに、検索対象の文から検

索キーの単語と同じ長さの単語を見出す候補単語検出手段と、

前記候補単語検出手段により見出された候補単語に対して、文字認識を行なって検索キーの単語か否かの判断を行なう最終文字認識手段とを有していることを特徴とするデータ検索装置。

【請求項25】 図形的な特徴や使用される文字に一定の規則性が見出される定型的な記載型式のデータを検出するデータ検出方法であって、

10 検出対象のデータの定型性をあらかじめ入力する定型情報入力ステップと、

上記あらかじめ入力された定型情報の中から第一に図形的な特徴を利用して候補データを検出する图形利用候補データ検出手段と、

前記图形利用候補データ検出手段にて検出されたデータを対象として文字認識を行なうことにより更に候補データを絞り込む文字認識利用候補データ検出手段とを有していることを特徴とするデータ検出方法。

【請求項26】 図形的な特徴や使用される文字に一定の規則性が見出される定型的な記載型式のデータを検出するデータ検出装置であって、

20 検出対象のデータの定型性をあらかじめ入力する定型情報入力手段と、

上記あらかじめ入力された定型情報の中から第一に図形的な特徴を利用して候補データを検出する图形利用候補データ検出手段と、

前記图形利用候補データ検出手段にて検出されたデータを対象として文字認識を行なうことにより更に候補データを絞り込む文字認識利用候補データ検出手段とを有していることを特徴とするデータ検出装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明は、データ検索に関し、特に計算機システムや情報システムにおける各種形式、量のデータを検索する方法及び装置に関する。

【0002】

【従来の技術】 近年、電子化された大量のデータを速やかに表示、参照することを可能とすべくデータ検索装置が広く利用されている。ところで、このようなデータ検索装置は、全文検索を行なう場合もあるが、一般的には大量のデータを取り扱うため、検索に際してキーワードやデータファイルの作成日の範囲の指定等の条件（検索用の鍵）を与え、これにより大まかな絞り込みを行い、その結果としての候補データ内容やそのファイルの表題のみや第一頁のみを表示させるのが普通である。

40 【0003】 以下、このようなデータ検索装置を採用したデータ保存装置の一例を図を参照しつつ説明する。図18は、その検索装置を含むデータ保存装置の構成図である。本図に示すように、この装置は、読込部1と、保

存部2と、検索指示部3と、検索部4と、出力部7とを

備えている。なお、実際には、以上その他、一旦検索した候補データ等を一時記憶する一時記憶部、表示等された候補データのタイトル等から意図するものを指定する指定部等を有するが、これらは本発明の趣旨には直接的には関係せず、また周知技術かつ自明の事項であるため、その記載等は省略する。そして、これは後の実施例でも同様である。

【0004】読み込部1は、文書を文字、表、グラフ、図形等のコードの集合として読み込んだり、あるいはL/E(光/電気)変換による読み取りスキャナを用いたり、ネットワークを介したりして読み込んだりする。保存部2は、読み込部1により読み込まれたデータを、データの管理者、入力者により別途作成されたキーワードや入力日等と対応づけて保存する。

【0005】検索指示部3は、検索者の操作により入力された検索の起動や停止、各種検索条件などのコマンドに応じた指示を検索部4に通知する。検索部4は、検索指示部3からの指示に基づいて、検索条件に適合するデータを候補データとして保存部2から順に読み出す。出力部7は、検索部4により読み出された候補データをあらかじめのプログラムや検索者の指示により定められた出力形式で出力する。

【0006】検索者は、この出力結果を視認して本来必要としている目的のデータであるか否かを確認する。そして、所望のデータが得られない場合には新たなキーワードを作成する等して再度の検索が指示されることとなる。なおここで、出力結果は、出力データの表題、データの属性(種別)、その他ケースによってはデータ内容そのものやその一部等であり、出力形式は、CRTへの表示、用紙への印刷、スピーカへの音声等である。

【0007】次に、以上は主に文書、画像等の保存されているデータの検索であったが、同様の検索、その結果の表示等は各種検索条件の指定の一環としてチャネルの選択等があるものの、文字放送等においてもなされる。同じく、マルチメディアシステムの一環としての、ビデオ・オン・デマンドシステムにおけるサービスに供されている映像情報の選択等においてもなされる。

【0008】

【発明が解決しようとする課題】しかしながら、上記従来の構成では、以下のような問題点がある。第1に、一般にデータ量が多いこともあり、検索用鍵としての検索条件の選定作業が繁雑である。次に、データの登録者が検索者の便宜のため登録にあたりあらかじめ作成してあるキーワードと、検索者が検索のため入力するキーワードとが必ずしも一致するとは限らない。そして、若し両方のキーワードが極力一致するようにその作成と入力の手順を厳格に定めたならば、逆にキーワードの作成と入力のための処理が複雑、煩雑となる。このため、この手順を統一することによる効果にも自ずと限界がある。結果的に、検索の対象となるデータが膨大な場合には、ど

のように検索鍵を設定しても検索結果として出力される候補データも大量となるが、この出力や検索者による確認も手間と時間がかかる。

【0009】更に、全文検索がなされることも少なからずあるが、この場合には該当する候補データ量がどうしても多くなりがちであり、ひいてはこれらの不都合も増大しがちである。第2に、上記第1の課題に付随した面もあるが、保存されているデータや目的とするデータの内容如何によっては、検索条件を変更しての再検索が必要となることが多くなり、応答性が損なわれる。具体的には、保存されているデータが階層的な場合、同じく膨大な場合、同じく記載内容の趣旨と検索者の参照せんとする意図とが相違する場合等である。

【0010】第3に、検索の対象とするデータが過去の新聞記事等極めて膨大な一方で型式が一定の場合には、検索の対象となるデータを検索者が一度でも視認したり、読み込んだことがあるときには、その際の視覚的な記憶が上記型式と照らしあわせると有力な検索の鍵となり得るが、それに対する考慮がなされていない。更には、形式が一定であることが予想される場合も同様に検索に利用しうるが、それへの考慮がなされていない。

【0011】第4に、上記第1の課題に共通した面もあるが、検索された候補データが大量の場合にその全ての印刷出力が指示されているようなときには、出力そのものに時間と費用がかかる等し、この一方で課題のみ出力させようとなれば同一課題の候補データが多数存在する等し、このため適切な出力形式の選択も必要となる。第5に、検索装置は、出力の如何にかかわらず、検出した候補データを一時記憶しておく必要が多いが、検索された候補データが大量の場合には、そのために必要な記憶資源の量も膨大となる。ひいては、装置も高価となる。

【0012】第6に、上記第3の課題と共通する面もあるが、そして理工学の分野等では特にそうであるが、图形的な特徴が検索に利用されていない。第7に、データの利用頻度等が反映されていない。そして、これらの問題点は、検索の対象となるデータが単なる文献等ではなく、文字放送、通信回線を介しての有料の映画、音楽、その他の情報の提供等のときにも生じる。

【0013】本発明は、かかる課題に鑑みてされたものであり、検索対象のデータの量や種類にかかわらず効率的なデータ検索方法及び装置を提供することを目的とする。同じく、検索対象のデータの量が膨大なときに、経費、手間、時間等の面から効率的なデータ検索方法及び装置を提供することを目的とする。同じく、検索対象のデータの量、種類、型式や利用度等があらかじめ判明しており、特に検索者が目的とするデータの量、種類、型式等につき何等かの記憶を有しているときやデータの性質からその量、種類、形式等が予想されるとき等のときに、効率的なデータ検索方法及び装置を提供することを目的とする。

【0014】同じく、検索対象のデータが、単なる文献等でなく、文字放送、通信回線を介しての映画、音楽、その他の情報の提供等のときにも同様な効果が得られる効率的なデータ検索方法及び装置を提供することを目的とする。

【0015】

【課題を解決するための手段】上記目的を達成するため、請求項1の発明は、文字、表、グラフ、図形等で構成されるデータ、コードデータやビットデータ等を検索し、結果をデータファイル名、データ内容、データファイル属性、データ特徴（その他放映）等の形式で出力するデータ検索方法であって、検索者の指示により、保存されている、文字放送されている、あるいはビデオ・オン・デマンド等で利用者のサービスに提供等されている文献、映像、画像、音楽等のデータから候補データを所定の手順で読み出す検索ステップと、前記検索ステップにおいて読み出された全データのチャネル数、個数等の情報量を測定する情報量測定ステップと、前記情報量測定ステップによる測定値が所定の値を越えた場合には、所定値未満の場合とは別形式の出力を行う、検索者に出力方法の選択指示をあおぐ、検索者に他の検索内容による再検索を要請する少なくも一を行う検索結果出力適切対応ステップとを有していることを特徴としている。

【0016】請求項2の発明は、文字、表、グラフ、図形等で構成されるデータを検索し、結果をデータファイル名、データ内容、データファイル属性、データ特徴などの形式で出力するデータ検索装置であって、検索者の指示により、保存、放送、提供等されているデータから候補データを読み出す検索手段と、前記検索手段において読み出された全データのビット数等の情報量を測定する情報量測定手段と、前記情報量測定手段による測定結果が所定値を越えた場合には、所定値未満の場合と別形式の出力を行う、検索者に出力方法の選択指示をあおぐ、検索者に他のキーワード等の入力等の他の検索内容による再検索を要請するの少なくも一を行う検索結果出力適切対応手段とを有していることを特徴としている。

【0017】請求項3の発明は、文字、表、グラフ、図形等で構成されるデータを検索し、結果をデータファイル名、データ内容、データファイル属性、データ特徴などの形式で出力するデータ検索方法であって、検索者の指示により、保存、放送、提供等されているデータから候補データを読み出す検索ステップと、前記検索ステップにおいて読み出された全データの情報量を測定する情報量測定ステップと、前記情報量測定ステップの測定値があらかじめ決められた制限値を超えているか否かを判定し、制限値を超えている場合には、前記検索ステップにより読み出されたデータを、あらかじめ入力される等した所定の規則により分類する検索データ分類ステップと、前記検索データ分類ステップにて分類された各分類毎に出力を行うか否かを所定の手順で判断する出力デー

タ分類選択ステップとを有していることを特徴としている。

【0018】請求項4の発明は、文字、表、グラフ、図形等で構成されるデータを検索し、結果をデータファイル名、データ内容、データファイル属性、データ特徴などの形式で出力するデータ検索装置であって、検索者の指示により、保存、放送、提供等されているデータから候補データを読み出す検索手段と、前記検索手段において読み出された全データの情報量を計測する情報量測定手段と、前記情報量測定手段の測定値があらかじめ決められた制限値を超えているか否かを判定し、制限値を超えている場合には、前記検索手段により読み出されたデータを、別途入力等されている所定の規則により分類する（例えば、一般向け外国語文献専門書等）検索データ分類手段と、前記検索データ分類手段にて分類された各分類毎に、出力を行うか否かを所定の手順で判断する（例えば、外国語文献と専門書は出力しない。一般向けは100頁以下のもののみ出力する等。）出力データ分類選択手段とを有していることを特徴としている。

【0019】請求項5の発明は、文字、表、グラフ、図形等で構成されるデータを検索し、結果をデータファイル名、データ内容、データファイル属性、データ特徴などの形式で出力するデータ検索方法であって、検索者の指示により、保存、放送、提供等されているデータから候補データを読み出す検索ステップと、前記検索ステップにおいて読み出された全データの情報量を測定する情報量測定ステップと、前記情報量測定ステップの測定値があらかじめ決められた制限値を超えているか否かを判定して、制限値を超えている場合には、前記検索ステップにより読み出されたデータについて所定のデータファイル属性あるいはデータ特徴（この他、言語、利用度）等により分類する検索データ分類ステップと、前記検索データ分類ステップにて分類された各分類毎に、同じ分類のデータが複数存在することとなる分類のデータのみ出力するよう制御する同一分類データ出力制御ステップとを有していることを特徴としている。

【0020】請求項6の発明は、文字、表、グラフ、図形等で構成されるデータを検索し、結果をデータファイル名、データ内容、データファイル属性、データ特徴などの形式で出力するデータ検索装置であって、検索者の指示により、保存、放送、提供等されているデータから候補データを読み出す検索手段と、前記検索手段において読み出された全データの情報量を測定する情報量測定手段と、前記情報量測定手段の測定値があらかじめ決められた制限値を超えているか否かを判定し、制限値を超えている場合には、前記検索手段により読み出されたデータについて特定のデータファイル属性あるいはデータ特徴等により分類する検索データ分類手段と、前記検索データ分類手段にて分類された各分類毎に、同じ分類のデータが複数存在することとなる分類のデータのみ出力

するよう制御する同一分類データ出力制御手段とを有していることを特徴としている。

【0021】請求項7の発明は、文字、表、グラフ、図形等で構成されるデータを検索し、結果をデータファイル名、データ内容、データファイル属性、データ特徴などの形式で出力するデータ検索方法であって、検索者の指示により、保存、放送、提供等されているデータから候補データを読み出す検索ステップと、前記検索ステップにて読み出された全データの情報量を測定する情報量測定ステップと、前記情報量測定ステップの測定値があらかじめ決められた制限値を超えているか否かを判定し、制限値を超えている場合には、前記検索ステップにより読み出されたデータについて所定のデータファイル属性あるいはデータ特徴等により分類する分類ステップと、前記分類ステップの結果、同じ分類のデータが存在しないデータのみ出力するよう制御する非同一属性データ出力制御ステップとを有していることを特徴としている。

【0022】請求項8の発明は、文字、表、グラフ、図形等で構成されるデータを検索し、結果をデータファイル名、データ内容、データファイル属性、データ特徴などの形式で出力するデータ検索装置であって、検索者の指示により、前記保存、放送、提供等されているデータから候補データを読み出す検索手段と、前記検索手段において読み出された全データの情報量を測定する情報量測定手段と、前記情報量測定手段の測定値があらかじめ決められた制限値を超えているか否かを判定し、制限値を超えている場合には、前記検索手段により読み出されたデータについて特定のデータファイル属性あるいはデータ特徴等により分類する検索データ分類手段と、前記検索データ分類手段にて分類された各分類毎に、同じ分類のデータが存在しないデータのみ出力するよう制御する非同一分類データ出力制御手段とを有していることを特徴としている。

【0023】請求項9の発明は、文字、表、グラフ、図形等で構成されるデータを検索し、結果をデータファイル名、データ内容、データファイル属性、データ特徴などの形式で出力するデータ検索方法であって、検索者の指示により、保存、放送、提供等されているデータから候補データを読み出す検索ステップと、前記検索ステップにおいて読み出された全データの情報量を測定する情報量測定ステップと、前記情報量測定ステップにより測定された情報量をあらかじめ定められたいずれの範囲内であるかを判定する範囲判定ステップと、前記範囲判定ステップの判定した範囲に応じて前記検索ステップから読み出されたデータの出力方法、出力内容の少くも一を制御する出力制御ステップとを有していることを特徴としている。

【0024】請求項10の発明は、文字、表、グラフ、図形等で構成されるデータを検索し、結果をデータファ

イル名、データ内容、データファイル属性、データ特徴などの形式で出力するデータ検索装置であって、検索者の指示により、保存、放送、提供等されているデータから候補データを読み出す検索手段と、前記検索手段において読み出された全データの情報量を測定する情報量測定手段と、前記情報量測定手段により測定された情報量の値があらかじめ決められた範囲であるかを判定する範囲判定手段と、前記範囲判定手段の判定した範囲に応じて前記検索手段から読み出されたデータの出力方法若しくは出力内容の少くも一（含む、両方）を制御する出力制御手段とを有していることを特徴としている。

【0025】請求項11の発明は、文字、表、グラフ、図形等で構成されるデータを検索し、結果をデータファイル名、データ内容、データファイル属性、データ特徴などの形式で出力するデータ検索方法であって、検索者の指示により、保存、放送、提供等されているデータから候補データを読み出す検索ステップと、前記検索ステップにおいて読み出された全データの情報量を測定する情報量測定ステップと、前記情報量測定ステップの測定値があらかじめ設定された制限値を超えているか否かを判定し、制限値を超えている場合には、前記検索ステップにより読み出されたデータをあらかじめ決められた規則に従って並び替えて出力するよう制御する出力順制御ステップとを有していることを特徴としている。

【0026】請求項12の発明は、文字、表、グラフ、図形等で構成されるデータを検索し、結果をデータファイル名、データ内容、データファイル属性、データ特徴などの形式で出力するデータ検索装置であって、検索者の指示により、保存、放送、提供等されているデータから候補データを読み出す検索手段と、前記検索手段において読み出された全データの情報量を計測し、その値があらかじめ設定された制限値を超えているか否かを判定する情報量測定手段と、前記情報量測定手段により測定された情報量が制限値を超えている場合、前記検索手段により読み出されたデータをあらかじめ決められた条件で並び替え、その順で出力するよう出力手段に指示する出力順制御手段とを有していることを特徴としている。

【0027】請求項13の発明は、文字、表、グラフ、図形等で構成されるデータを検索し、結果をデータファイル名、データ内容、データファイル属性、データ特徴などの形式で出力するデータ検索方法であって、検索者の指示により、保存、放送、提供等されているデータから候補データを読み出す検索ステップと、前記検索ステップにおいて読み出された各データの情報量を順に確認しつつ計測する情報量順計測ステップと、前記情報量順計測ステップの結果に基づきそれまでの総和があらかじめ定められた制限値を超えているか否かを判定する情報量測定ステップと、前記情報量測定ステップにより測定された情報量の総和が制限値を超えた場合には、制限値

に該当するデータ以降に読み出されたデータは出力しな

い若しくは制限値を越えた時点で検索を中止するよう制御する出力検索制限御ステップとを有していることを特徴としている。請求項14の発明は、文字、表、グラフ、図形等で構成されるデータを検索し、結果をデータファイル名、データ内容、データファイル属性、データ特徴などの形式で出力するデータ検索装置であって、使用者の操作に基づく検索起動指示により、保存、放送、提供等されているデータから候補データを読み出す検索手段と、前記検索手段において読み出された各データの情報量を順に測定する情報量測定手段と、前記情報量測定手段の測定した総和があらかじめ設定された制限値を超えていたか否かを判定し、制限値を超えた場合には、それ以降に前記検索手段により読み出されたデータは出力しない若しくは制限値を越えた時点で検索を中止するよう制御する出力検索制限手段とを有していることを特徴としている。

【0028】請求項15の発明においては、前記範囲判定ステップにて所定の範囲にあると判断されたならば、前記検索ステップにて検索された候補データの中に画像データが含まれているか否かを連続する黒画素の寸法や枠の有無等で判断する画像データ存在判断ステップを有し、前記範囲判定ステップは更に、同じく画像データが存在すると判断されたならば、その頁(枚)数等のデータ量を測定する画像データ量測定小ステップを有し、前記手段制御ステップは更に、前記画像データ存在判断ステップにて存在すると判断されたならば、画像データを先に出力されるよう候補データの配列を調整する画像データ優先配列小ステップと、前記画像データ量測定小ステップにて画像データ量が所定値以上と判断されたならば、画像データが出力される際にこれを縮尺して、同一表示面に複数の画像データが同時に表示されるよう制御する画像データ出力表示制御小ステップとを有していることを特徴としている。

【0029】請求項16の発明においては、前記範囲判定手段にて所定の範囲にあると判断されたならば、前記検索手段にて検索された候補データの中に画像データが含まれているか否かを判断する画像データ存在判断手段を有し、前記範囲判定手段は更に、同じく画像データが存在すると判断されたならば、その放送チャネル数等のデータ量を通信規則に基づく等して測定する画像データ量測定部を有し、前記出力制御手段は更に、前記画像データ存在判断にて存在すると判断されたならば、画像データを先に出力されるよう候補データの配列を調整する画像データ優先配列部と、前記画像データ量測定部にて画像データが所定値以上と判断されたならば、画像データが出力される際にこれを縮尺して、同一表示面に複数の画像データが同時に表示されるよう制御する画像データ出力表示制御部とを有していることを特徴としている。

【0030】請求項17の発明においては、検索装置に

10 よる検索の対象とされている、そして保存されているデータ若しくは保存するデータに経験則、データの意味内容、キーワード等から定められた本来の検索の目的とされる確率についての情報を検索に先立ってデータ保存者等が付加する利用度情報付加ステップを有し、前記出力制御ステップは、前記範囲判定ステップの判定した範囲に応じて前記検索ステップから読み出されたデータの出力方法若しくは出力内容を制御するに際して、前記利用度情報付加ステップにて付加された情報を読み出し、確率の低いデータは出力を中止する等して反映させる付加利用度情報反映出力制御小ステップを有していることを特徴としている。

【0031】請求項18の発明においては、検索装置による検索の対象とされている、そして保存されているデータ若しくは保存するデータに経験則、データの意味内容、キーワード等から定められた本来の検索の目的とされる確率についての情報を検索に先立って付加する利用度情報付加手段を有し、前記出力制御手段は、前記範囲判定手段の判定した範囲に応じて前記検索手段から読み出されたデータの出力方法若しくは出力内容を制御するに際して、前記利用度情報付加手段にて付加された情報を反映させる付加利用度情報反映出力制御部を有していることを特徴としている。

【0032】請求項19の発明は、前記出力制限御ステップにより出力を制限されなかったデータについては、あらかじめ定められた所定の規則に並び替えて出力するよう制御する出力順制御ステップを有していることを特徴としている。請求項20の発明は、前記出力制限手段により出力を制限されなかったデータについて、あらかじめ決められた条件で並び替え、その順で出力するよう出力手段に指示する出力順制御手段を有していることを特徴としている。

【0033】請求項21の発明においては、記事の意味内容から定まる記事の欄の分類と該記事の記載されている位置若しくはこれに併せての記載部分の形状、寸法、記載部やその近辺の所定の記号、図形、写真、枠等記載部分を特定するのに役立つ視覚的特徴との間に一定の関係の存在する日刊新聞紙、雑誌、データベース化された名刺等の定型文書から目的とする記載を検索する方法において、検索目的たる記事等が記載されている少くも一の種類の定型文書について、その定型性についての情報を検索に先立ってあらかじめ入力しておく定型性情報入力ステップと、検索目的たる記事の意味内容から定まる記事の欄の分類若しくはこれに併せての該記事の記載されている定型文書の種類を検索キーとして入力する欄定型情報入力ステップと、検索目的たる記事中の文章、単語、漢字等の文字の少くも一方について、少くも一個を検索キーとして入力する若しくはこれに併せて記載位置、近辺の所定の記号等をも検索キーとして入力する記事情報入力ステップとを有していることを特徴としている。

る。

【0034】請求項22の発明においては、記事の意味内容から定まる記事の欄の分類と該記事の記載されている位置若しくはこれに併せての記載部分の形状、寸法、記載部やその近辺の所定の記号、図形、写真、枠等記載部分を特定するのに役立つ視覚的特徴との間に一定の関係の存在する定型文書から目的とする記載を検索する装置において、検索目的たる記事等が記載されている少くも一の種類の定型文書について、その定型性についての情報を検索に先立ってあらかじめ入力され記憶しているHD、磁気ディスク等からなる定型性情報保持手段と、検索目的たる記事の意味内容から定まる記事の欄の分類若しくはこれに併せての該記事の記載されている定型文書の種類を検索キーとして入力するキーボード等を有する欄定型情報入力手段と、検索目的たる記事中の文章、単語、漢字やドイツ文字等の文字の少くも一について、少くも一個を検索キーとして入力する若しくはこれに併せて記載位置、近辺の所定の記号等をも検索キーとして入力する記事情報入力手段とを有していることを特徴としている。

【0035】請求項23の発明においては、相前後する単語間に余白を設ける言語で記載された文を対象として、文中に存在する単語を検索キーとして、所望するデータの検索を行なうデータ検索方法であって、検索対象の文中の単語間の余白から文中の各単語の長さを求める単語長検出ステップと、検索キーの単語長を所定の手順で求める検索キー単語長検出ステップと、前記両検出ステップの結果をもとに、検索対象の文中から検索キーの単語と同じ長さの単語を見出す候補単語検出ステップと、前記候補単語検出ステップにより見出された候補単語に対して、文字認識を行なって検索キーの単語か否かの判断を行なう最終文字認識ステップとを有していることを特徴としていて、請求項24の発明においては、相前後する単語間に余白を設ける言語で記載された文を対象として、文中に存在する単語を検索キーとして、所望するデータの検索を行なうデータ検索装置であって、検索対象の文中の単語間の余白から文中の各単語の長さを求める単語長検出手段と、検索キーの単語長を所定の手順で求める検索キー単語長検出手段と、前記両検出手段の結果をもとに、検索対象の文中から検索キーの単語と同じ長さの単語を見出す候補単語検出手段と、前記候補単語検出手段により見出された候補単語に対して、文字認識を行なって検索キーの単語か否かの判断を行なう最終文字認識手段とを有していることを特徴としていて、請求項25の発明においては、图形的な特徴や使用される文字に一定の規則性が検出される定型的な記載型式のデータを検出するデータ検出方法であって、検出対象のデータの定型性をあらかじめ入力する定型情報入力ステップと、上記あらかじめ入力された定型情報の中から第一に图形的な特徴を利用して候補データを見出

する图形利用候補データ検出ステップと、前記图形利用候補データ検出ステップにて検出されたデータを対象として文字認識を行なうことにより更に候補データを絞り込む文字認識利用候補データ検出ステップとを有していることを特徴としている。

【0036】請求項26の発明においては、图形的な特徴や使用される文字に一定の規則性が見出される定型的な記載型式のデータを検出するデータ検出装置であって、検出対象のデータの定型性をあらかじめ入力する定型情報入力手段と、上記あらかじめ入力された定型情報の中から第一に图形的な特徴を利用して候補データを検出する图形利用候補データ検出手段と、前記图形利用候補データ検出手段にて検出されたデータを対象として文字認識を行なうことにより更に候補データを絞り込む文字認識利用候補データ検出手段とを有していることを特徴としている。

【0037】

【作用】上記構成により、請求項1の発明においては、請求項1の発明においては、文字、表、グラフ、図形等で構成されるデータを検索し、結果をデータファイル名、データ内容、データファイル属性、データ特徴などの形式で印刷したりCRTに表示したりして出力するデータ検索方法であって、以下の作用がなされる。検索ステップにて、検索者（含む、その依頼者等）の指示により、保存、放送、提供等されているデータから候補データが読み出される。情報量測定ステップにて、前記検索ステップにおいて読み出された全データの情報量が測定される。検索結果出力適切対応ステップにて、前記情報量測定ステップによる測定値が所定の値を越えた場合には所定値未満の場合とは別形式の出力を行う、検索者に出力方法の選択指示をあおぐ、検索者に他の検索内容による再検索を要請する、の少なくも一がその旨をCRTに表示する等して行われる。

【0038】請求項2の発明においては、文字、表、グラフ、図形等で構成されるデータを検索し、結果をデータファイル名、データ内容、データファイル属性、データ特徴などの形式で出力するデータ検索装置であって、以下の作用がなされる。検索手段が、検索者の指示により、保存、放送、提供等されているデータから候補データを読み出す。情報量測定手段が、前記検索手段において読み出された全データの情報量を測定する。検索結果出力適切対応手段が、前記情報量測定手段による測定結果が所定値を越えた場合には、所定値未満の場合とは別形式の出力を行なう、検索者に出力方法の選択指示をあおぐ、検索者に他の検索内容による再検索を要請するの少なくも一を行う。

【0039】請求項3の発明においては、文字、表、グラフ、図形等で構成されるデータを検索し、結果をデータファイル名、データ内容、データファイル属性、データ特徴などの形式で出力するデータ検索方法であって、

以下の作用がなされる。検索ステップにて、検索者の指示により、保存、放送、提供等されているデータから候補データが読み出される。情報量測定ステップにて、前記検索ステップにおいて読み出された全データの情報量を測定される。検索データ分類ステップにて、前記情報量測定ステップの測定値があらかじめ決められた制限値を超えていたか否かを判定し、制限値を超えていた場合には、前記検索ステップにより読み出されたデータを所定の規則により分類される。出力データ分類選択ステップにて、前記検索データ分類ステップにて分類された各分類毎に出力を行うか否かが所定の手順で判断される。

【0040】請求項4の発明においては、文字、表、グラフ、図形等で構成されるデータを検索し、結果をデータファイル名、データ内容、データファイル属性、データ特徴などの形式で出力するデータ検索装置であって、以下の作用がなされる。検索手段が、検索者の指示により、保存、放送、提供等されているデータから候補データを読み出される。情報量測定手段が、前記検索手段において読み出された全データの情報量をページ数の総和を求める等して計測する。検索データ分類手段が、前記情報量測定手段の測定値があらかじめ決められた制限値を超えていたか否かを判定し、制限値を超えていた場合には、前記検索手段により読み出されたデータを所定の規則により分類する。出力データ分類選択手段が、前記検索データ分類手段にて分類された各分類毎に、出力を行うか否かを所定の手順で判断する。

【0041】請求項5の発明においては、文字、表、グラフ、図形等で構成されるデータを検索し、結果をデータファイル名、データ内容、データファイル属性、データ特徴などの形式で出力するデータ検索方法であって、以下の作用がなされる。検索ステップにて、検索者の指示により、保存、放送、提供等されているデータから候補データが読み出される。情報量測定ステップにて、前記検索ステップにおいて読み出された全データの情報量が測定される。検索データ分類ステップにて、前記情報量測定ステップの測定値があらかじめ決められた制限値を超えていたか否かが判定され、制限値を超えていた場合には、前記検索ステップにより読み出されたデータについてあらかじめ付加されていたり、電波に通信規則

(規約)等にもとづき付加されていたりする所定のデータファイル属性あるいはデータ特徴や所定の制御信号等により分類される。同一分類データ出力制御ステップにて、前記検索データ分類ステップにて分類された各分類毎に、同じ分類のデータが複数存在することとなる分類のデータのみ、例えば日本語文字放送のみ出力するよう制御される。

【0042】請求項6の発明においては、文字、表、グラフ、図形等で構成されるデータを検索し、結果をデータファイル名、データ内容、データファイル属性、データ特徴などの形式で出力するデータ検索装置であって、

以下の作用がなされる。検索手段が、検索者の指示により、保存、放送、提供等されているデータから候補データを読み出される。情報量測定手段が、前記検索手段において読み出された全データの情報量を測定する。検索データ分類手段が、前記情報量測定手段の測定値があらかじめ決められた制限値を超えていたか否かを判定し、制限値を超えていた場合には、前記検索手段により読み出されたデータについて特定のデータファイル属性あるいはデータ特徴等により分類する。同一分類データ出力制御手段が、前記検索データ分類手段にて分類された各分類毎に、同じ分類のデータが複数存在することとなる分類のデータのみ出力するよう制御する。

【0043】請求項7の発明においては、文字、表、グラフ、図形等で構成されるデータを検索し、結果をデータファイル名、データ内容、データファイル属性、データ特徴などの形式で出力するデータ検索方法であって、以下の作用がなされる。検索ステップにて、検索者の指示により、保存、放送、提供等されているデータから候補データが読み出される。情報量測定ステップにて、前記検索ステップにて読み出された全データの情報量が測定される。分類ステップにて、前記情報量測定ステップの測定値があらかじめ決められた制限値を超えていたか否かが判定され、制限値を超えていた場合には、前記検索ステップにより読み出されたデータについて所定のデータファイル属性あるいはデータ特徴等により分類される。非同一属性データ出力制御ステップにて、前記分類ステップの結果、同じ分類のデータが存在しないデータ、例えば外国語データのみ出力するよう制御される。

【0044】請求項8の発明においては、文字、表、グラフ、図形等で構成されるデータを検索し、結果をデータファイル名、データ内容、データファイル属性、データ特徴などの形式で出力するデータ検索装置であって、以下の作用がなされる。検索手段が、検索者の指示により、前記保存、放送、提供等されているデータから候補データを読み出される。情報量測定手段が、前記検索手段において読み出された全データの情報量を測定する。検索データ分類手段が、前記情報量測定手段の測定値があらかじめ決められた制限値を超えていたか否かを判定し、制限値を超えていた場合には、前記検索手段により読み出されたデータについて特定のデータファイル属性あるいはデータ特徴等により分類する。非同一分類データ出力制御手段が、前記検索データ分類手段にて分類された各分類毎に、同じ分類のデータが存在しないデータのみ出力するよう制御する。

【0045】請求項9の発明においては、文字、表、グラフ、図形等で構成されるデータを検索し、結果をデータファイル名、データ内容、データファイル属性、データ特徴などの形式で出力するデータ検索方法であって、以下の作用がなされる。検索ステップにて、検索者の指示により、保存、放送、提供等されているデータから候

補データが読み出される。情報量測定ステップにて、前記検索ステップにおいて読み出された全データの情報量が測定される。範囲判定ステップにて、前記情報量測定ステップにより測定された情報量があらかじめ定められたいずれの範囲内であるかを判定する。出力制御ステップにて、前記範囲判定ステップの判定した範囲に応じて前記検索ステップから読み出されたデータの出力方法や内容が制御される。

【0046】請求項10の発明においては、文字、表、グラフ、図形等で構成されるデータを検索し、結果をデータファイル名、データ内容、データファイル属性、データ特徴などの形式で出力するデータ検索装置であつて、以下の作用がなされる。検索手段が、検索者の指示により、保存、放送、提供等されているデータから候補データを読み出す。情報量測定手段が、前記検索手段において読み出された全データの情報量を測定する。範囲判定部が、前記情報量測定手段により測定された情報量の値があらかじめ決められたどの範囲であるかを判定する。出力制御手段が、前記範囲判定ステップの判定した範囲に応じて前記検索手段から読み出されたデータの出力方法若しくは出力内容を制御する。

【0047】請求項11の発明においては、文字、表、グラフ、図形等で構成されるデータを検索し、結果をデータファイル名、データ内容、データファイル属性、データ特徴などの形式で出力するデータ検索方法であつて、以下の作用がなされる。検索ステップにて、検索者の指示により、保存、放送、提供等されているデータから候補データが読み出される。情報量測定ステップにて、前記検索ステップにおいて読み出された全データの情報量を測定される。出力順制御ステップにて、前記情報量測定ステップの測定値があらかじめ設定された制限値を超えているか否かを判定し、制限値を超えている場合には、前記検索ステップにより読み出されたデータがあらかじめ決められた規則に従って並び替えて出力するよう制御する。

【0048】請求項12の発明においては、文字、表、グラフ、図形等で構成されるデータを検索し、結果をデータファイル名、データ内容、データファイル属性、データ特徴などの形式で出力するデータ検索装置であつて、以下の作用がなされる。検索手段が、検索者の指示により、保存、放送、提供等されているデータから候補データを読み出す。情報量測定手段が、前記検索手段において読み出された全データの情報量を計測し、その値があらかじめ設定された制限値を超えているか否かを判定する。出力順制御手段が、前記情報量測定手段により測定された情報量が制限値を超えている場合、前記検索手段により読み出されたデータをあらかじめ決められた条件で並び替え、その順で出力するよう出力手段に指示する。

【0049】請求項13の発明においては、文字、表、

グラフ、図形等で構成されるデータを検索し、結果をデータファイル名、データ内容、データファイル属性、データ特徴などの形式で出力するデータ検索方法であつて、以下の作用がなされる。検索ステップにて、検索者の指示により、保存、放送、提供等されているデータから候補データが読み出される。情報量順計測ステップにて、前記検索ステップにおいて読み出された各データの情報量を順に確認しつつ計測する。情報量測定ステップにて、前記情報量順計測ステップの結果に基づきそれまでの総和があらかじめ定められた制限値を超えているか否かを判定する。出力検索制限御ステップにて、前記情報量測定ステップにより測定された情報量の総和が制限値を超えた場合には、制限値に該当するデータ以降に読み出されたデータは出力しない若しくは制限値を越えた時点で検索を中止するよう制御する。

【0050】請求項14の発明においては、文字、表、グラフ、図形等で構成されるデータを検索し、結果をデータファイル名、データ内容、データファイル属性、データ特徴などの形式で出力するデータ検索装置であつて、以下の作用がなされる。検索手段が、使用者の操作に基づく検索起動指示により、保存、放送、提供等されているデータから候補データを読み出す。情報量測定手段が、前記検索手段において読み出された各データの情報量を順に測定する。出力検索制限手段が、前記情報量測定手段の測定した総和があらかじめ設定された制限値を超えているか否かを判定し、制限値を超えた場合には、それ以降に前記検索手段により読み出されたデータは出力しない若しくは制限値を越えた時点で検索を中止するよう制御する。

【0051】請求項15の発明においては画像データ存在判断ステップにて、前記範囲判定ステップにて所定の範囲にあると判断されたならば、前記検索ステップにて検索された候補データの中に画像データが含まれているか否かを判断する。前記範囲判定ステップ内の、画像データ量測定小ステップにて、同じく画像データが存在すると判断されたならば、そのデータ量を測定する。

【0052】前記出力制御ステップ内の、画像データ優先配列小ステップにて、前記画像データ存在判断ステップにて存在すると判断されたならば、画像データを先に出力されるよう候補データの配列を調整する。同じく、画像データ出力表示制御小ステップにて、前記画像データ量測定小ステップにて画像データ量が所定値以上と判断されたならば、画像データが出力される際にこれを縮尺して、同一表示面に複数の画像データが同時に出力されるよう制御する。

【0053】請求項17の発明においては、利用度情報付加ステップにて、検索装置による検索の対象とされている、そして保存されているデータ若しくは保存するデータに経験則、データの意味内容、キーワード等から定められた本来の検索の目的とされる確率についての情報

を検索に先立って付加する。前記出力制御ステップ内の、付加利用度情報反映出力制御小ステップにて、前記範囲判定ステップの判定した範囲に応じて前記検索ステップから読み出されたデータの出力方法若しくは出力内容を制御するに際して、前記利用度情報付加ステップにて付加された情報を反映させる。

【0054】請求項19の発明においては、出力順制御ステップが、前記出力検索制限御ステップにより出力を制限されないこととなつたデータについては、あらかじめ定められた所定の規則にのつとつて並び替えて出力するよう制御する。請求項16、18、20の発明においては、各請求項15、17、19の発明と同じ作用がなされる。

【0055】請求項21の発明においては、記事の意味内容から定まる記事の欄の分類と該記事の記載されている位置若しくはこれに併せての記載部分の形状、寸法、記載部やその近辺の所定の記号、図形、写真、枠等記載部分を特定するのに役立つ視覚的特徴との間に一定の関係の存在する定型文書から目的とする記載を検索する方法において、以下の作用がなされる。

【0056】定型性情報入力ステップにて、データの保存者等より検索目的たる記事等が記載されている少くも一の種類の定型文書について、その定型性についての情報を検索に先立つてあらかじめ入力しておく。欄定型情報入力ステップにて、検索者より検索目的たる記事の意味内容から定まる記事の欄の分類若しくはこれに併せての該記事の記載されている定型文書の種類を検索キーとして入力される。

【0057】記事情報入力ステップにて、検索目的たる記事中の文章、単語、文字の少くも一について、少くも一個を検索キーとして入力する若しくはこれに併せて記載位置、近辺の所定の記号等をも検索キーとして入力される。請求項22の発明においては、記事の意味内容から定まる記事の欄の分類と該記事の記載されている位置若しくはこれに併せての記載部分の形状、寸法、記載部やその近辺の所定の記号、図形、写真、枠等記載部分を特定するのに役立つ視覚的特徴との間に一定の関係の存在する定型文書から目的とする記載を検索する装置において、以下の作用がなされる。

【0058】定型性情報保持手段が、検索目的たる記事等が記載されている少くも一の種類の定型文書について、その定型性についての情報を検索に先立つてデータ保存者等よりあらかじめ入力され記憶している。欄定型情報入力手段が、検索目的たる記事の意味内容から定まる記事の欄の分類若しくはこれに併せての該記事の記載されている定型文書の種類を検索キーとして入力する。

【0059】記事情報入力手段が、検索目的たる記事中の文章、単語、漢字の少くも一方について、少くも一個を検索キーとして入力する若しくはこれに併せて記載位置、近辺の所定の記号その他色彩等をも検索キーとして

入力する。請求項23の発明においては、相前後する単語間に余白を設ける言語で記載された文を対象として、文中に存在する単語を検索キーとして、所望するデータの検索を行なうデータ検索にて、以下の作用がされる単語長検出ステップにて検索対象の文中の単語間の余白から文中の各単語の長さを求める。検索キー単語長検出ステップにて、検索キーの単語長を所定の手順で求める。

【0060】候補単語検出ステップにて、前記両検出ステップの結果をもとに、検索対象の文中から検索キーの単語と同じ長さの単語を見出す。最終文字認識ステップにて、前記候補単語検出ステップにより見出された候補単語に対して、文字認識を行なつて検索キーの単語か否かの判断を行なう。請求項24の発明においては、請求項23の発明と同じ作用がなされる。

【0061】請求項25の発明においては、図形的な特徴や使用される文字に一定の規則性が検出される定型的な記載型式のデータを検出するデータ検出方法にて、以下の作用がなされる。定型情報入力ステップにて、検出対象のデータの定型性をあらかじめ入力する。図形利用候補データ検出ステップにて、上記あらかじめ入力された定型情報の中から第一に図形的な特徴を利用して候補データを検出する。文字認識利用候補データ検出ステップにて、前記図形利用候補データ検出ステップにて検出されたデータを対象として文字認識を行なうことにより更に候補データを絞り込む。

【0062】請求項26の発明においては、請求項25の発明と同じ作用がなされる。

【0063】
30 【実施例】以下、本発明を実施例に基づいて説明する。
(第1実施例) 図1の(1)に、本発明に係るデータ検索装置の第1実施例を探り入れたデータ保存装置の構成を示す。本図に示すように、このデータ保存装置は、読み込部1と、保存部2と、検索部3と、検索指示部4と、情報量測定部5と、検索出力適切対応部6と、出力部7とを備えている。そして、情報量測定部5と検索出力適切対応部6とを除くこれら各部は、基本的には従来のものと同一である。このため同一の符号を付してある。以下、従来の技術欄と重複する部分もあるが、これら各部を説明する。

【0064】読み込部1は、キーボード、L/E変換部、レーザディスクや磁気ディスク等からの読み取り部等を有し、文字、表、グラフ、図形などからなるデータをコードの集合やビットマップデータとして直接的、間接的に読み込む。保存部2は、磁気ディスク等からなり、読み込部1により読み込まれた本来のデータを別途若しくは読み込みの際にデータ作成者や装置の管理者等により作成された検索用のキーワードや入力日時やデータの標題等と対応づけて記憶し、管理し、また保存する。そして、この対応付けにより、検索用のキーワード、入力日

時、データの標題等からデータやデータの標題を検索することが可能となっている。そして、勿論、全文検索も可能としている。

【0065】更に、検索結果の出力の便宜のために各データ毎にその標題、データがビットデータかコードデータか等のデータの属性、データの内容や分類、記憶されている始点と終点の番地 (address) 等がいわばタイトルとして併せて記憶され、またこのタイトルは独立して読み出し可能とされ、更に標題等は表示等可能とされている。この様子を図2に示す。また、これにより、タイトルから該当するデータを検索することも可能である。なお、このことは図書館における書籍の分類、収納状態や管理、検索の手段と基本的には同じである。このため、その詳細な説明は省略する。

【0066】検索指示部3は、検索者のキーワードの入力等の操作に基づいて、保存部2に保存されているデータの検索を検索部4に指示する。検索部4は、検索指示部3からの指示に基づいて、該当するデータやそのタイトルのみをキーワード等を鍵として保存部2から検出す。情報量測定部5は、検索部4により検出された結果たる候補データの情報量、本実施例では、個数を測定し、その測定値を検索出力適切対応部6に出力する。この際の手順を図1の(2)に示す。

【0067】検索出力適切対応部6は、処理対応表61と処理対応表作成部62とを有している。そして、処理対応表61は、図3の(1)に示すように、検索の結果得られた候補データが全て検出された場合には、その情報量の値に応じて、出力方法の指示や再検索の指示を行うべく、処理内容が登録されている。なお、この表の内容は、処理、対応表作成部62によりあらかじめ入力されている。この際、情報量に応じての指示内容は、検索者が1度に評価、確認、参照等可能な量に収まるようにされているが、基準となるデータ数そのものについては、データ管理者の経験に基づいて決定されている。具体的には、ある検索条件での検索結果のデータ数が10未満の場合には、各データの内容まで出力し、10以上50未満の場合はデータの標題(タイトル又はファイル名)のみ表示する。更に、データ数が50以上の場合には、検索条件を変更して再検索するよう検索者に要請するようにしている。

【0068】以上のもとで、検索出力適切対応部6は、情報量測定部5から通知されたデータの個数をこの処理対応表の値と比較して、適切な出力を行なうこととなる。すなわち、データ数が10未満ならば、検出された全ての候補データの内容を読み出して表示する。10以上50未満ならばタイトルのみ表示す。50以上ならば、検索された候補データの個数と併せて別途のキーワードの入力等、検索条件を変更して再検索する旨検索者に表示する。この際の、出力の選択の手順を図3の(2)に示す。

【0069】なお、検索の結果得られた候補データの標題のみを一旦出力するよう別途あらかじめ検索者から指示されている場合には、たとえデータ数が10未満であっても、標題のみが出力されるのは勿論である。次に、上記装置の検索動作について、図4に示すフローチャートを参照しながら説明する。

【0070】検索に先立ち、あらかじめ読込部1から文字、表、グラフ、図形などの電子化されたコードの集合で構成されるデータが読み込まれ、保存部2に保存されている。検索者が、検索指示部3に検索鍵の入力を含む起動を指示すると、検索指示部3がその指示を検索部4に通知する(ステップS1)。

【0071】これにより検索部4が、保存部2から検索条件に該当するデータを1つずつ検出し、そのタイトルを情報量測定部5に送る(ステップS2)。検索条件に合致する次のデータがあるか否かを判断し(ステップS3)、次データがあればステップS2に戻って次のデータを読み出し、検索対象の保存されているデータがなくなるまでステップS2～S3の動作が繰り返される。

【0072】検索条件に合致するデータの検出が全て終了した場合は、情報量測定部5が求められた検索結果の候補データのタイトル数を求める(ステップS4)。そして、その値から、処理対応表を参照し、再検索が必要か否かを判断する。候補データが全く存在しなかつた場合や検索結果が多くて別のキーワードの入力等再検索が必要な場合には、その旨検索者に表示し、併せてステップS2に戻って他の検索条件の入力を待ち、これによる再検索を行うこととなる。一方、再検索の必要がなければ、検索の結果のデータの出力形式を、処理対応表に適合するように選択する(ステップS5)。なお、本ステップの具体的な内容は図3の(2)に示す。

【0073】このため、候補データのタイトル数が10以上50未満ならば、情報量測定部5に通知されている全候補データのタイトルが出力され、同じく10未満ならばそのタイトルに相当するデータが保存部2から全て読み出され、出力されることとなる。検索結果のデータの所定の出力を終了すると、他の検索者等による次の検索の入力を待つ(ステップS7)。

【0074】なお、本実施例では、検索結果のデータの情報量としてデータ個数を採用しているが、これは図書館類似のデータベースにおいて、検索者や候補データが多数存在する場合の応答性を考慮したものである。従って、応答性を特に問題としないならば、検索結果たる候補データのページ総数であってもよい。その他、データをディスプレイに表示した時の表示領域サイズ、行数、あるいはビット数等のサイズでもよい。

【0075】更に、検索対象のデータの種別に応じて、適宜文字数、グラフの数、画像やイメージの数等の(情報)量を基に出力を変更するようにしてもよい。なお、これらの場合には、情報量測定部5には、各候補データ

タのページ数、行数、データの属性等をそのタイトル等から調べたり、また各候補データのタイトルにページ数やデータ属性等についての情報が付されていない場合には、直接保存部2から読み出して調べたりする候補データ量内容調査部分を有しているのは勿論である。

【0076】また、情報量としてデータの行数やビット数が採用された場合には、画像データや映像データに対しては、必要に応じて「画像1枚は文書20行に相当」等の換算基準も入力されたりするのも勿論である。次に、情報量によるデータ処理の指示の内容やその判断基準となる値は、本装置の管理者があらかじめの経験に基づいて最適な値を設定しておいてもよいし、検索者が検索の対象とするデータの内容に応じて最適なものを指定するようにしても良い。

【0077】同じく、再検索時の条件については、検索者に新たな検索条件を参考に示してその中から指定せることによっても良い。具体的には、前回の検索時に与えた条件にキーワードを追加する形式を薦めても良いし、画像情報化してページめぐり態様で検索する等全く新たな検索条件を薦めるようにしても良い。そして、これらの推奨される形式は、本装置の管理者等がその経験に基づいてあらかじめ本装置に入力しておき、必要に応じてCRTに表示される等してもよい。

【0078】さらにまた、候補データの出力に際しての処理とは、検索結果を印刷出力するよう検索者に指示されている場合において、その頁数が極めて多いときは、検索者の指示にもかかわらず、頁数が極めて多い旨とタイトルのみ印刷出力する等であってもよい。以上の説明でわかるように、本実施例では、検索結果たる候補データの量によってその出力を変更等するので、候補データが多量の場合に、検索効率の向上、出力に伴う費用の低減等が図れる。

【0079】またこの際、検索者やデータの保管者の経験等も有効に活用しえるので、上記効果が一層増す。なお、これら第1実施例の変形例は、後に説明する他の多くの実施例においても適用しえるのは勿論である。

(第2実施例) 図5の(1)に、本発明に係るデータ検索装置の第2実施例を取り入れたデータ保存装置の構成図を示す。このデータ保存装置は、読み込部1と、保存部2と、検索部3と、検索指示部4と、情報量測定部5と、出力データ選択部6と、出力部7とを備えている。これら各部のうち、先の第1実施例と同一の構成のものについては同一の符号を付すことによりその説明は省略し、以下固有の構成たる出力データ選択部26を中心にして説明する。

【0080】出力データ選択部26は、情報量測定部5により測定された情報量が制限値を超えている場合には、検索部4により読みだされたデータについて特定の条件により分類し、この上で分類内容毎に出力を行うか否かを指示する。出力部7は、出力データ選択部26に

より出力を指定されたデータを出力する。

【0081】次にこの動作の内容をより詳細、具体的に説明する。情報量測定部25は、検索条件に合致する全てのデータのタイトルを得、その値があらかじめ決められた制限値を超えた場合には、出力データ選択部26にその旨を通知する。このため、たとえば、制限値としての情報量がデータ数30とした場合に、検索条件に合致した候補データ数が25ならば、出力部7に全データを全て出力するよう指示する。ところが、もし候補データ数が40だった場合には、情報量測定部5は出力データ選択部26に制限値を超えていることを通知する。この通知を受けた出力データ選択部26は、特定の条件によりデータを分類し、この分類毎に出力するか否かを指示することとなる。その条件であるが、本実施例では、

(1) ビットマップデータであれコードデータであれ画像のあるデータを優先して表示する。

【0082】(2) 画像データが多数存在するならばCRTの表示面を4分割し、画像データも1/4に縮小して一の表示面に同時に4つの画像データが表示されるようする。

(3) 文字データであるならば、ページ数の少ないものを優先して表示する。

(4) 一応は、全てのデータを表示する。

【0083】という4種のものを採用している。以上の結果、候補データから検索者が、いわゆるパラパラとページをめくる態様で、すなわち視覚による画像認識で目的とするデータの検索を行なうこととなる。以上説明した、本実施例のデータ出力の選択手順を図5の(2)に示す。

【0084】なお、それでも目的とするデータが得られない場合には、他の手段によることとなるのは勿論である。また、画像データの1/4の縮小は、上下方向は走査線を一本毎に間引き、左右方向は画素を一個毎に間引くこと等によりなされるが、その詳細は周知技術であるため省略する。また、候補データの再配列等も同様に省略する。

【0085】以上のごとく、本実施例では確認容易な候補データから調べていくこととなるため、迅速な検索が可能となる。次に、本実施例の変形例について説明する。

(第2実施例、第1変形例) 一度でも見たことのある資料ならば、その資料の表紙の図形的な印象や内部の絵画等は人の記憶に残っていることが多い。そこで、候補データが、多数存在し、しかもそれらが書籍等多数のページからなるものであるときには、図形としての表紙を、もし該当する記載の有る一部のページに絵画等が含まれている場合には、そのページをも併せて確認のためCRTに表示可能なようにしている。

【0086】これにより、ページめぐり態様で図形化された候補データを視覚的に確認することにより迅速に確

認することが可能となる。

(第2実施例、第2変形例) 本実施例は、文字放送の中から意図する放送を迅速に選択することに関する。なお、本実施例の構成そのものは、文書データの読み込部1、保存部2に換えて文字放送局があり、検索指示部3や検索部4に換えてチャネル切換部やチャネル条件入力部が存在するだけであるため、わざわざ固有の構成を図示しての説明は省略する。

【0087】さて、文字放送の対象となっているデータの内容によっては、視覚的かつ同一分類ではあるが内容が多種のもののが多々存在する。例えば、天気予報等である。具体的には、本日の午前の予報、午後の予報、明日の予報、月間予報、長期予報、それも北海道、関東、九州等地域毎になされている。この場合、現在東京に住んでおり明日北海道へ出張する人にとって、九州の予報等全く必要がない。

【0088】このため、本変形例では天気予報が検索対象とされた場合に、多数の内容の天気予報が候補データとして検索され、この上で縮尺されて同一画面に同時に表示されることになる。次いで、視覚により意図するものを見出した視聴者は、縮尺画像に同時に表示されているチャネル番号をもとにその放送局にチャネルをあわせる。これにより、意図する明日の北海道の予報、若しこれがないならば明日の東日本、北日本の予報を見ることとなる。

【0089】この結果、一々ページをめくる態様での検索は不要となる。似たことは、株式の放送等についてもいえる。ひいては、東京株式市場やニューヨーク株式市場での、自分の注目する会社の株価等の確認が迅速になされる。

(第3実施例) 図6の(1)に、本発明に係るデータ検索装置の第3実施例を探り入れたデータ保存装置の構成を示す。このデータ保存装置は、読み込部1と、保存部2と、検索部3と、検索指示部4と、情報量測定部5と、同一分類データ抽出部36と、出力部7とを備えている。これら各部のうち、先の第1実施例と同じ構成ものについては、同一符号を付すことによりその説明は省略し、以下固有の構成たる同一データ属性抽出部36を中心にして説明する。

【0090】情報量測定部5は、検索部4により読み出されたデータの情報量を測定し、その値があらかじめ決められた制限値を超えているか否かを判定する。同一分類データ抽出部36は、情報量測定部5により測定された情報量が制限値を超えている場合には、検索部4により読みだされた候補データを特定のデータファイル属性あるいはデータ特徴等により分類し、同じ分類のデータが複数存在することとなる分類に属するデータのみを抽出して出力するよう出力部7に指示する。このため、かかる場合には、出力部7は、同一分類データ抽出部36により出力を指定されたデータを出力することとなる。

【0091】次に、この動作の内容をより詳細、具体的に説明する。たとえば、情報量をデータ数とし、あらかじめ決められた制限値が30の場合に、検索条件に合致するデータ数が25のとき、出力部7はデータを全て出力する。ところが、合致するデータ数が32のときには、情報量測定部5は同一分類データ抽出部36に制限値を超えていることを通知する。この通知を受けた同一分類データ抽出部36は、あらかじめ与えられた条件で分類したときに同じ属性を持つデータが複数個存在するデータについてのみ出力するよう出力部7に指示する。

例えば、今日作成されたデータという条件で検索し、図7に示すように39個が該当したとする。

【0092】「データの内容」で分類した結果、天気予報、会議議事録、テストサンプルデータがそれぞれ複数存在するため、これらのデータ分類に属するデータのみが表示されることになる。以上のデータ出力の選択の手順を図6の(2)に示す。これにより、検索対象の可能性の高い候補データ、利用に供される頻度が高いデータが優先して出力されるため、検索効率が向上する。

【0093】また、文字放送においては、言語による分類を採用していた場合に、外国語による放送の除外等をなしえる。

(第4実施例) 図8に、本発明に係るデータ検索装置の第4実施例を採用したデータ保存装置の構成を示す。このデータ保存装置は、読み込部1と、保存部2と、検索部3と、検索指示部4と、情報量測定部5と、非同一分類データ抽出部46と、出力部7とを備えている。これら各部のうち、先の第1実施例と同じ構成ものについては、同一の符号を付すことにより、その説明は省略し、本実施例固有の構成を中心にして説明する。

【0094】情報量測定部5は、検索部4により読み出されたデータの情報量を測定し、その値があらかじめ決められた制限値を超えているか否かを判定する。非同一分類データ抽出部46は、情報量測定部5により読みだされたデータを特定のデータファイルあるいはデータ特徴等により分類し、同じ分類となるデータが存在しない分類データのみを抽出して出力するよう出力部7に指示する。このため出力部7は、非同一分類データ抽出部46により出力を指定されたデータのみを出力することとなる。

【0095】次にその動作の内容をより具体的に説明する。情報量測定部5は、検索された各データの情報量を測定し、検索条件に合致する全てのデータ情報量を得、その値があらかじめ決められた制限値を超えた場合は、非同一分類データ抽出部46に通知する。このため、たとえば、情報量をデータ数とした場合に、検索条件に合致するデータ数が25のとき、あらかじめ決められた制限値が30の場合には、出力部7はデータを全て出力することとなる。

【0096】ところが、合致するデータ数が39のとき

には、情報量測定部5は非同一分類データ抽出部4.6に制限値を超えてることを通知する。この通知を受けた非同一属性データ抽出部4.6は、あらかじめ与えられた条件で分類したときに同じ分類に属するデータが他に存在しない、すなわち1の分類に1のデータが存在することとなるデータについてのみ出力するよう出力部7に指示する。

【0097】例えば、分類の条件をデータの内容とし、図7に示すような候補データが検索された場合には、日報、仕様書がそれぞれ1つだけ存在するため、これら2データが出力されることになる。ここに、データの内容であるが、本実施例では、図2に示すタイトルの一部たる分類欄の一番下の段の記載を利用している。

【0098】これにより、内容が一般的であるため多数の分類に属することとなるデータの出力が排除される。ひいては、専門的なデータを検索する場合に、検索条件に最っとも適合しているデータのみ出力され、データの検索効率が向上する。また、在日かつ日本語に上手でない外国人が文字放送を見ようとする場合、日本語放送は排除されたりもする。

【0099】なお、候補データを分類する条件は、データの保管管理者により一括して指定されてもよいし、また検索者により検索の都度指定されるようにしてもよい。また、この際、候補データのキーワードに着目するようにしてもよい。すなわち、ある特定の候補データをとってきたとき、その候補データに付されているキーワードをキーワードとする他の候補データが所定数存在するならば、その特定の候補データは出力しないというふうにしてもよい。これにより、一般的な候補データの排除がより達成される。

(第3実施例と第4実施例の折衷例)

(1) 各分類毎に、その分類に属するデータから、所定数づつ出力するようにしている。これにより、候補データの内容の巾が広がり、一般利用者による検索に対して、効率的となる。

(2) 複数の分類に属することとなるデータを優先して出力するようにしている。これにより、候補データの内容の巾が広がり、一般利用者の検索の効率が増大する。

【第5実施例】図9に、本発明に係るデータ検索装置の第5実施例を採用したデータ保存装置の構成を示す。このデータ保存装置は、読み込部1と、保存部2と、検索部3と、検索指示部4と、情報量測定部5と、出力内容制御部5.6と、出力部7とを備えている。これら各部のうち、先の第1実施例と同じ構成のものについては、同一の符号を付すことによりその説明は省略し、本実施例固有の構成を中心にして説明する。

【0100】情報量測定部5は、検索部4により読み出された全候補データの情報量を測定する。出力内容制御部5.6は、図10に示すような出力方法対応表を保持し

ており、情報量測定部5により測定された情報量の値からデータの出力方法を出力方法対応表を用いて求め、出力部7に指示する。このため、出力部7は、出力内容制御部5.6により指定された出力方法でデータを出力する。

【0101】次に、この動作の内容をより具体的に説明する。情報量測定部5は、検索された各候補データの情報量としてのタイトル数を測定し、その総和を求めて出力内容制御部5.6に通知する。出力内容制御部5.6は、情報量測定部5により測定された情報量の値と出力方法対応表から、検索部4より読みだされた候補データの出力方法を出力部7に指示する。このため、たとえば、情報量をデータ数とした場合に検索条件に合致するデータ数が25のとき、図10に示す出力方法対応表から、出力部7にはデータタイトルとデータファイルの属性を出力するよう指示がなされることになる。

【0102】以上の他、検索者は候補データ数が多い場合には、例えば画像を含む候補データのみ表示させて、図形としての候補データを視覚的に確認し、意図するものを発見し、また意図せざるものを排除等することも可能である。ひいては、検索効率が向上する。

【第6実施例】図11に、本発明に係るデータ検索装置の第6実施例を採用したデータ保存装置の構成を示す。このデータ保存装置は、読み込部1と、保存部2と、検索部3と、検索指示部4と、情報量測定部5と、出力順制御部6.6と、出力部7とを備えている。そして、先の第1実施例と同じ構成のものには同一の符号を付してその説明は省略し、本実施例固有の構成を中心にして説明する。

【0103】情報量測定部5は、検索部4により読み出された全データの情報量を測定し、その値があらかじめ決められた制限値を超えているか否かを判定する。出力順制御部6.6は、情報量測定部5により測定された情報量の値が制限値を超えている場合、検索部4より読みだされた候補データをあらかじめ決められた条件で並び替え、その順で出力するよう出力部7に指示する。このため、出力部7は、出力順制御部6.6により指定された出力順で候補データを出力することとなる。

【0104】次に、この動作の内容をより具体的に説明する。情報量測定部5は各候補データの情報量を測定し、その総和があらかじめ決められた制限値を超えた場合には出力順制御部6.6にその旨通知する。この通知を受けた出力順制御部6.6は、検索部4より読みだされたデータをあらかじめ決められた条件、例えば作成日付順に並びかえ、この上で出力部7に送る。一方、情報量が制限値を超えた場合には、検索部4により読み出された順で出力することとなる。

【0105】次に、作成日付順で出力することであるが、これは最近、近年の資料であるならば検索者の印象、記憶が鮮明であるだけに、多くのキーワード等を入

力することにより候補データを絞り込むことが可能である。しかし、それができず、候補データが多数検索されたのであるから、本来意図するデータは昔のものである可能性が高いことを考慮したものである。

【0106】次に、何度も改定がなされている書籍については、昔の版が検索対象となることは少ない。また、内容によっては、昔のデータが検索対象とれることは少ない。このため、データの内容、分類に応じて作成日時や入力された日時が新しい順に配列されて出力されるようにもよい。これにより、例えば新聞社における記事作成等において、読者、記者の印象に残っている事件ほど引例として使用され、過去の事件は引例として使用されることは少なくなり、ひいては検索結果の利用性の向上も図られる。

【0107】また、ビデオ・オン・デマンドシステムにおいて、フィルムの指定に先立ち、原作者がキーワードとして入力された場合、大作家の名作ほど何度も映画化されていることが多いが、この場合にはその製作年代の新しいもののタイトルが表示されることとなる。そして、これにより、検索効率が向上する。

(第7実施例) 図12に、本発明に係るデータ検索装置の第7実施例を採用したデータ保存装置の構成を示す。このデータ保存装置は、読み込部1と、保存部2と、検索部3と、検索指示部4と、情報量測定部5と、出力制限部7と、出力部7とを備えている。なお、先の第1実施例と同じ構成のものについては、同じ符号を付すことによりその説明を省略し、本実施例固有の構成を中心に説明する。

【0108】出力制限部7と、情報量測定部5により測定された情報量の総和が制限値を超えた場合には、それ以降に検索部4より読みだされたデータを出力しないよう出力部7に指示する。このため、出力部7は、出力制限部7と、情報量測定部5により出力を制限されなかつたデータのみを出力することとなる。次に、この動作の内容を具体的に説明する。

【0109】情報量測定部5は、検索された各データの情報量を測定し、その総和があらかじめ決められた制限値を超えた時点で出力制限部7と通知する。この通知を受けた出力制限部7は、検索部4から読み出されたデータを順に出力部7に転送するのを中止する。このため、例えば、制限値がデータ数で20とあらかじめ決められていた場合は、検索部4から読み出された順に20個のデータのみが出力部7に出力されることとなる。

【0110】これにより、大図書館等で検索装置を利用する人が多い場合に、装置を多数の人が利用しうる。ことに、印字機の能力の都合で出力が、ひいては利用者数が制限される等の不都合がなくなる。また、候補データ数が20ともなれば以降の検索を中止して出力するため、装置の応答性も改善される。

【0111】なお、ここに制限値を20としたが、これ 50

はデータの属性、内容等に応じて適宜最適なものとするようにしていてもよい。また、本実施例では、データの保存は検索の対象とされる頻度の高い順に配列しておくのがより効率的となるであろう。

(第5実施例と第6実施例の変形例)

(1) 候補データの個数が制限値を越えた場合には、ページ数若しくは行数の少ないデータを優先して制限値だけ出力するようとしている。

【0112】これにより、新聞、雑誌等広い範囲の候補データが入手しえる。

(2) 上記(1)とは逆に、ページ数若しくは行数の多いデータのみ制限値の1/5の値のデータ数だけ優先して出力する。これにより、候補データの重複した記載を排除しえ、ひいては真に得たい資料のみ入手しえる確率が向上する。

(第8実施例) 図13に、本発明に係るデータ検索装置の第8実施例を採用したデータ保存装置の構成を示す。このデータ保存装置は、読み込部1と、保存部2と、検索部3と、検索指示部4と、情報量測定部5と、出力制限部8と、出力順制御部8と、出力部8とを備えている。そして、先の第1実施例と同一の構成のものについては同一の符号を付すことにより、その説明は省略し、本実施例固有の構成を中心説明する。

【0113】情報量測定部5は、検索部4により読み出された各データの情報量を測定し、その総和があらかじめ決められた制限値を超えているか否かを判定する。出力制限部8と、情報量測定部5により測定された情報量の総和が制限値を超えた場合、それ以降に検索部4より読みだされたデータを出力しないよう出力順制御部8と指示する。

【0114】出力順制御部8は、出力制限部8により出力を制限されなかつたデータについて、あらかじめ決められた条件で並び替え、その順で出力するよう出力部8に指示する。このため、出力部8は、出力順制御部8により指示された順でデータを出力することとなる。次に、この動作の内容を具体的に説明する。

【0115】情報量測定部5は、検索された各データの情報量としてのタイトルを数え、その総和があらかじめ決められた制限値である30を超えた時点で出力制限部8と出力順制御部8にその旨を通知する。この通知を受けた出力制限部8は、これ以降検索部3から読み出されたデータを順に出力順制御部8に転送するのを行わない。

【0116】出力順制御部8は、この通知を受けた場合には、出力制限部8より連送されたデータをあらかじめ決められた条件、例えば作成日付順でソートし、出力部8に送る。このように、情報量が多い場合は先に読みだされたデータのみを出力させるため、使用者による検索結果の確認が容易となる。

【0117】また、結果出力の応答が早いため、効率的

なデータ検索を行うことができる。

(第9実施例) 本実施例は、検索の対象となっている資料、データ側にも工夫を凝らしてある。以下、最初にこの背景となっている事実について説明する。

【0118】図書館等で、閲覧の対象、貸出しの対象となっている書籍、資料は自ずと一定の傾向、性質がある。ひいては、検索の対象、目的となるデータについても一定の傾向、性質がある。具体的には、例えば古代日本の倭女王「卑弥呼」に関する資料を検索しようとする場合、通常の図書館で一般人が参照しようとする資料は、歴史学者等が一般人向け、あるいはアマチュア

(?) 向けに著した本等である。そして「三国志の魏志倭人伝」や「魏略の逸文」の本文そのもの(原文)が対象となることはまずない。

【0119】同じく、一般相対性理論についても大学の物理の教授等が一般人向けに解説した本が参考の対象となり、「HARMANN WEYL著「RAUM ZEITMATERIE」等が参考の対象となることはまずない。大学の図書館等では、逆であろう。そして、このことは大なり小なり他種のデータの検索に際しも見受けられる事実である。このため、検索者の如何によるが、滅多に検索の対象とされることのないデータには、あらかじめその旨とその理由を示す印が付されて保存、管理されている。

【0120】一般向けの図書館であるならば、その理由とは、

- (1) 専門書
- (2) 外国語(特に英語でない)。
- (3) 旧版等データそのものが古い。
- (4) 上記(1)、(2)とも重複するが難解である。
- (5) 狂信家の著書等内容に問題点が多い。
- (6) 訴訟記録等内容が特殊である。

(7) エロ、グロ等公序良識に反しかねないあるいは成人向けであり、一般的の利用に共するのに難がある。

【0121】等である。この一方で、よく利用される。ひいては検索の対象となる資料も大体の傾向が存在する。具体的には、(1)有名な先生、その道の大家の著した本、(2)官公庁による一般公衆等を対象とした公的な出版物や報告書、若しくはこれに準じた資料。

【0122】(3)全国的な日刊新聞紙、(4)著名な総合雑誌、(5)特定の出版社から刊行された本や雑誌、等である。このため、保存されているデータにはその旨の情報があらかじめ付されている。

【0123】図14に本実施例における資料の保存、管理の様子を示す。図2に比較した場合、タイトル欄の先に利用度とその理由を記載した欄が設けられているのが異なる。本図において、「工業所有権法逐条解説(第12版)」という本は、「官公庁による公的な出版物や報告書若しくはこれも準じた資料である」ため「利用者が大」であるのがわかる。

【0124】なお、この一方で第11版以前のもの(図示せず)は、「データそのものが古い。」という理由のもと「(3)」の符号を付された上で、その利用度は「小」となっているのは勿論である。次に、これらの資料が検索の候補データとなった場合等においては、この「利用度及びその理由」が出力の有無等の判断資料とされる。

【0125】従って、今まで説明してきた第1実施例から第8実施例までのデータ検索装置において、入力されたキーワード等の検索条件に適合する候補データ数が多量の場合には、利用頻度の高い資料のみ出力されるようになる。また、利用度の低い資料、そして多くの場合保存されている資料の過半を占める、は通常の検索ではそもそも候補データとされない。これにより、検索効率の向上を図る。

【0126】なお、この場合、滅多に利用されないため原則として検索対象とされない、そして出力をされないこととなる資料は、その様な資料が存在すること及び出力されない理由が表示等され、別途の入力等で出力されるようになっているのは勿論である。また、直接その資料を特定する入力がなされた場合等は、出力されるのも勿論である。

【0127】また、利用度は検索者の如何によって変化するため、これをも反映するようにもよい。この場合には、検索者が一般大衆、大学の先生、学生等のいずれに属するかも入力され、更に利用度のリストもそれらに応じて作成されているのは勿論である。以上の動作流れを図15に示す。

(第9実施例の変形例)

30 (1) 各資料毎に経験則等に基づく利用頻度を反映した重み付けをなしておき、候補データの量に応じて、利用頻度の高い資料より出力等されるようしている。

【0128】同じく、キーワードに対応しての利用頻度を調べ、これを反映させるようしている。具体的には、外国語、専門語がキーワードとして入力された場合には、専門書が高くされ、一般大衆向けの雑誌等は低く評価される等である。なお、この場合には、専門語にはこれを示す印が付されているのは勿論である。

(2) 外国語言語をも採用したビデオ、オン・デマンドシステムや国際的な通信網を介してのビデオ、オン・デマンドシステム(1の映像に対して、複数の言語の音声が対応する。)において、我国からの要求に際しては、映像に対応する音声は日本語とする。

(3) 外国語言語による文字放送において、我国での要求に対しては、日本語放送を優先する。

(第10実施例) 本実施例は、検索そのものに関する。以下、本実施例そのものの説明に先立ち、背景技術や背景事実について説明する。

【0129】特許情報、新聞、雑誌等、検索の対象として利用されることの多い書籍データ類は、記載内容と記

載位置や図形的特徴に一定の規則が存在する。新聞紙を例にとるならば、

- (1) 日付は、最上段の線の上に記載されている。
- (2) 第一面は社会、政治記事等、紙面(頁)と記事の意味内容が定められている。
- (3) 漫画は、4コマのものは最後から2面の最左側の上部に、政治、社会を風刺したものは第一頁めの面若しくはその裏面の下半分の位置に掲載される。また、枠でかこまれている。更に、4コマ漫画は最上段に標題が記載され、登場人物の会話(台詞)は線で囲まれている。
- (4) 通常の計報は、最後から2頁めの面の下側、かつ死亡者名の右側に黒線が付される。また、広く知られた人物では、同時に写真も掲載される。
- (5) 通常の人物写真的形状は長(正)方形又は橢(円)である。
- (6) 広告は、全面広告や広告欄を除き原則として紙面の最下部である。
- (7) 過半の記事は、直線等で囲まれたりして区分されている。

(8) 「天声人語」等所定の記載は、所定の位置に記載されるだけでなく、所定の図形が付されたりしている。

(9) 見出し等は、大きな文字とされる。等である。

【0130】ところで、現在手書き文字や印刷文字の機械装置による光学的読み速度は、ワークステーション等小型のものならば手書き文字であるならば10文/秒、多少大型のものならば100字/秒程度である。また、大企業等が備えるような本式なものならば、更にこの速度は大である。次に、印刷文字であるならば、新聞等字体、文字形状、寸法等が一定かつ主として当用漢字が使用されるものならばより早く、10万字/秒を超えるものさえ出現しつつある。また、コードデータならば、この速度は更に早い。

【0131】さて、実生活において、昔読んだことのある記事を再度読みたい、あるいは昔どこかで読んだことのある内容であるが再度あたってみたいと思うようなことが多々ある。特許関係における先行技術の調査、新聞社における記者が記事を書くに際しての調査等では、この要請は更に大きい。ところで、現在、新聞紙一面あたりの文字量は1万字である。この際、参照を欲する記事の内容が、広告、科学技術、家庭、政治等定まれば、これにより、該記載が新聞に記載されている紙面が自ずと定まり、更に随筆、直線で囲まれている等他種の情報も得られることが多い。従って、キーワード等の検索キーと併せて大まかな掲載年月日等検索者の何等かの記憶による検索候補データの絞り込みをなしえれば、後は定型性を利用して最終的な絞り込みを行ない、最後に参照を欲する記事中の文書中の記者の記憶に残っている単語や文字、特に漢字、をもとに当該記事の検索を行なうことが可能となる。この際、検索者の記憶に残っている単語や漢字は、その性質上新聞紙の同一面の他の記事に使用

されないものであることが多い。特に、複数の場合にはなおさらである。また、多くは見出し等大きな文字で記載されたものである。このため、これらをも検索に利用しえれば、更に容易となる。また、図形的な記憶があれば、これをも入力可能として、検索に利用しうるようすれば、更に効率は向上する。

【0132】本実施例は、これを利用したものである。以下、本実施例そのものを説明する。図16の(1)は、その構成図である。本図において161はデータ保存部であり、162は定型性情報保持部であり、163は一般検索部であり、164は定型性利用単語文字検索部であり、165は出力部であり、167は一般検索鍵入力部であり、168は単語文字入力部であり、169は欄定型情報入力部である。

【0133】定型性情報保存部162は、データの保存管理者が保存すべきデータをデータ保存部161に入力するに際して、その定型性情報を入力され、これを利用可能な態様で記憶している。一般検索部163は検索者が一般検索鍵入力部167から入力した、著者、記事の分類、月日、キーワード等の鍵をもとにデータ保存部161の保存されているデータを検索し、該当するものを定型性利用単語文字検索部164に出力する。単語文字入力部168は、検索者から記事中の単語、文字等最終的な検索鍵として入力されて欄定型情報入力部167は、検索目的の記事が有する各種特徴のうち定型性情報保持部162に保存されている定型性に関する特徴を第2段階の検索鍵として入力される。このため、検索者は入力に先立ち、必要な定型性情報についての指示書等を参照したりする。なお、ここに入力される検索鍵の内容は、新聞名、検索目的の記事の種類、例えば計報、家庭欄、料理欄、スポーツ、科学技術等の種類、その他記憶している内容で指示書に記載されている事項、例えば枠で囲まれていたか否か、記事中の写真や色彩の有無、字体、大きな見出しの有無、大体の掲載位置や記事の所属する欄の名称、例えば「余録」、「天声人語」等である。定型性利用単語文字検索部164は、一般検索部163から送られてきた候補データを欄定型情報入力部169から入力された定型性に関する情報を基に、定型性情報保持部162から必要な情報を読み出して絞り込みを行なう。この上で、単語文字入力部168から入力された記事中の単語や文字をもとに最終的な検索を行なう。これにより、目的とする記載が検索される。なお、本が検索目的である場合には、その本の厚さ、大きさ等の図形的特徴も定型性情報として使用される。そして、この場合には、保存部中の候補データの記憶されている場所(スペース)の広さとの比較も本の厚さによる絞り込みの一環としてなされる。

【0134】以上の検索の手順を図16の(2)に示す。なお、新聞、雑誌等の定型性は発行年等により適宜変更されるため、定型性情報保持部162の保持する定

型性の規則は年代等に応じて整理されてのは勿論である。

(第10実施例の第1変形例) 第10実施例は、日本語の新聞であったが、欧米系の言語の新聞にあっては、イタリック体、太字体等が定型性情報、検索キーとして採用され、あるいは入力されたりする。また、単語間に余白があるため、各単語や文字の認識に先立って、

「i」、「1」等単語長に影響を及ぼす文字、数字が認識対象の単語に含まれている場合には特に、単語長を基に図形的に先行検索を行ってもよい。また、ドイツ語にあっては、たとえ文中であっても名詞の先頭の文字は大文字とされる。従って、たとえ同じ文字数でも、名詞は動詞よりも単語長が長くなる。従って、これも各文字の認識に基づく単語認識に先立って、文中の候補単語の絞り込みに利用しうる。図17に、これらの様子及び手順を示す。本図において、(1)は英語文書の一部であり、(2)はドイツ語文であり、(3)は検索手順図である。本図の(1)において、同じ2文字からなる英単語であっても、「is」は「i」の存在のため「be」より単語長(黒画素の水平方向長さ)が短くなっているのがわかる。同様に、(2)において4文字からなる単語「Baum」は大文字「B」のため「hoch」より長くなっている。従って、文字認識に先立ち、単語長による絞り込みが可能となる。

(第10実施例の第2変形例) 次に、科学技術では特に、材料毎の温度と引張り強度の関係、同じく温度による腐蝕量等、テーブルやグラフで表現されるデータが検索対象とされることが多い。そしてまた、これらのデータは検索対象の文献にテーブルやグラフの形式で記載されていることが多い。

【0135】ところで、テーブルやグラフは、言語、技術分野を問わずほぼ同一の形式である。すなわち、まず第一に横書きであり、グラフならば二次元のXY座標で示され、かつX軸(左右方向)は時間等であり、またY軸は伸び等であったりする。テーブルならば、各材料毎に100℃、500℃等区切りのよい温度毎に比重や粘度が記載され、かつ所定の区分け線が上下、水平方向に記載されている等である。また、時間や物性値等は、例えば時間ならば、「時間」、「T」、「t」、「分」、「sec」等一定の形式で記載されている。このため、検索装置にテーブルやグラフの図形としての定型性や使用される文字の規則をあらかじめ記憶させておき、その記憶をもとに所望するテーブルやグラフを検出することが可能となる。なお、この場合にも、別途のキーワードによる文献の絞り込み、文字認識による最終確認等が併用されれば、その効率が一層高まるのは勿論である。そしてこの場合、図16の(1)に示す構成図は、定型性情報保持部162等換えて、テーブル、グラフに対する図形及び文字情報保持部等を有することとなる。また、図16の(2)に示す「欄定型情報による絞り込み」

は、「テーブル、グラフに対する図形的及び文字的に特徴づける情報による検出や絞り込み」となり、更に検索対象の文献は、新聞でなく、広く科学技術文献一般となる。これにより、難解な外国語候補データの各単語や文字の確認の手間もはぶける。

【0136】以上、本発明を実施例に基づいて説明してきたが、本発明は何も上記実施例に限定されないのは勿論である。すなわち、以下のようなものも本発明に含まれる。

10 (1) 製造等の都合で、本発明の一つの必要不可欠な構成要素(要件、ステップ、手段)を複数のものとしている、あるいは逆に複数の構成要素を一体のものとしている、更にはこれらを適宜組み合わせている。

(2) 現在の検索装置に改造を施したり、必要なプログラムを新規に記憶させたりして本発明に係わる作用、効果を発揮させる。

(3) 検索対象は、ビットマップデータ、コードデータ、文字放送に限定されず、ビデオ、オン・デマンドレスシステムにおける映画や音楽の放送としている。な

20 お、映画の場合、タイトルとしては標題の他に監督や主演俳優名等が含まれるのは勿論である。これにより、「伊豆の踊り子」、「青い山脈」、「侍日本」等何度も製作された映画の区別がなされる。同じく、交響曲ならば指揮者名も含まれる。

(5) また、各検索者は所持する個人用の、そして個人の作成したデータ保存装置であり、検索対象のデータは各人がワードプロセッサで作成した文である。

(6) 候補データが多量の場合には、特定の外国語、特定の著者、新聞記事文書

30 のみのデータ等所定の分類に属するものを排除するための新規な排除キー依頼が表示されるようになっている。

(7) 検索者の指示のもと、全文検索が可能となる。

【0137】

【発明の効果】以上説明してきたように、本発明によれば、入力されたデータを保存する保存手段と、検索された全候補データの情報量を計測し、その値により通常の出力や検索者の指示する出力そのままにするか否かを判断したり、出力方法を選定したり、再検索等を検索者にあおいだりするので、検索結果の確認が容易となる。

40 【0138】また、検索された全候補データの情報量があらかじめ決められた制限値を超えているか否かを判定し、制限値を超えている場合には、特定の条件により検索結果を分類し、出力を行うか否か等を定めたりするので、検索結果のデータの情報量を適切な値とするよう制御可能となり、検索結果の確認が容易となり、検索の効率が向上する。

【0139】また、同じく検索された全候補データの情報量が制限値を超えている場合には、検索されたデータについて特定のデータファイル属性あるいはデータ特徴等により分類し、同じ属性のデータが複数存在する分類

のデータのみを抽出して出力するため、検索結果情報量が適切な値となり、検索目的のデータの種類等によっては、検索者による確認が容易となる。

【0140】また、同じく検索された全候補データの情報量が制限値を超えている場合には、検索されたデータについて特定のデータファイル属性あるいはデータ特徴等により分類し、同じ属性のデータが存在しない分類となるデータのみを抽出して出力するので、検索結果のデータの情報量が適切な値になり、検索目的のデータの種類等によっては、検索者による確認が容易となる。

【0141】また、同じく検索された全候補データの情報量が制限値を超えている場合には、検索されたデータをあらかじめ決められた重要度等の条件で並び替えて出力するので、検索者による結果の確認や整理、利用が容易となる。また、同じく検索された各候補データの情報量が制限値を超えた場合には、それ以降検索されたデータは出力を中止したり検索そのものを中止したりするので、検索者による結果の確認が容易かつ応答性の向上が図れる。また、多数の検索者の利用に便となる。また、出力を制限されなかったデータについては、あらかじめ決められた条件で並び替えて、その順で出力するので、先に読みだされたデータのみを出力させることとなり、検索者による結果の確認が容易となり、かつ応答性が向上する。

【0142】また、検索手段された全データの情報量があらかじめ決められたどの範囲であるかを判断し、その判断結果に応じて、出力方法を変更するので、検索者による結果の確認が容易となる。また、文字放送等視覚化されたデータ、画像データ等は、縮尺した上同時に複数表示するので、検索者が本来必要としているデータの確認や入手が迅速となる。

【0143】また、検索された候補データが多量の場合で、一方このことを知らない検索者による印刷出力が指示されているようなときには、その旨の注意喚起等なされるため、無駄な出費の削減が図れる。また、滅多に利用されないデータ、逆によく利用されるデータ等を区分けして検索するので、検索効率が向上する。

【0144】更に、定型文書中のデータの検索に、検索者の記憶するデータ中の単語、漢字等の機械的（含む、ソフト的）読み取りと文書の定型性の图形的特徴による絞り込みを利用するので、検索効率が向上する。更にまた、グラフ、テーブル等言語、記載文献の種類の如何を問わず定型性を有するデータの検索が可能となる。

【図面の簡単な説明】

【図1】本発明に係るデータ検索装置の第1実施例を採用したデータ保存装置の構成とデータ数測定の手順を示す図である。

【図2】上記実施例におけるデータの保存、管理の様子を概念的に示した図である。

【図3】上記実施例における処理対応表の内容の一例と

その表に基づく出力の選択の手順を示す図である。

【図4】上記実施例におけるデータ検索の動作流れ図である。

【図5】本発明に係るデータ検索装置の第2実施例を採用したデータ保存装置の構成図である。

【図6】本発明に係るデータ検索装置の第3実施例を採用したデータ保存装置の構成とデータ出力の選択の手順を示す図である。

【図7】第3及び第4実施例におけるデータ検索部の出力内容の一例を示す図である。

【図8】本発明に係るデータ検索装置の第4実施例を採用したデータ保存装置の構成図である。

【図9】本発明に係るデータ検索装置の第5実施例を採用したデータ保存装置の構成図である。

【図10】第5実施例における出力方法対応表の内容の一例である。

【図11】本発明に係るデータ検索装置の第6実施例を採用したデータ保存装置の構成図である。

【図12】本発明に係るデータ検索装置の第7実施例を採用したデータ保存装置の構成図である。

【図13】本発明に係るデータ検索装置の第8実施例を採用したデータ保存装置の構成図である。

【図14】本発明の第9実施例におけるデータの保存、管理の様子を概念的に示した図である。

【図15】本発明の第9実施例の動作流れ図である。

【図16】本発明の第10実施例の構成と手順を示す図である。

【図17】本発明の第10実施例の第1変形例の原理、背景の事実と処理手順を示した図である。

【図18】従来の文書検索装置を採用したデータ保存装置の構成図である。

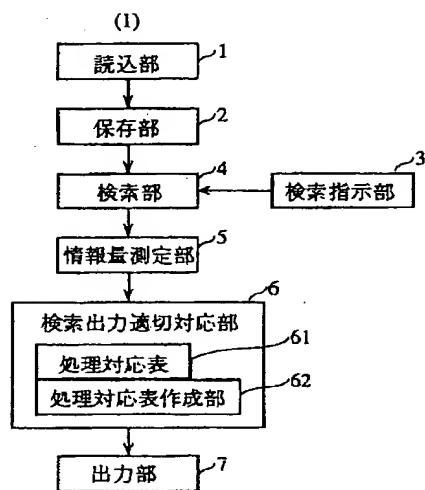
【符号の説明】

1	読み込部
2	保存部
3	検索指示部
4	検索部
5	情報量測定部
6	検索出力適切対応部
6 1	処理対応表
6 2	処理対応表作成部
7	出力部
2 6	出力データ選択部
3 6	同一分類データ抽出部
4 6	非同一分類データ抽出部
5 6	出力内容制御部
6 6	出力順制御部
7 6	出力制限部
8 6	出力制御部
8 7	出力順制御部
1 6 1	データ保存部

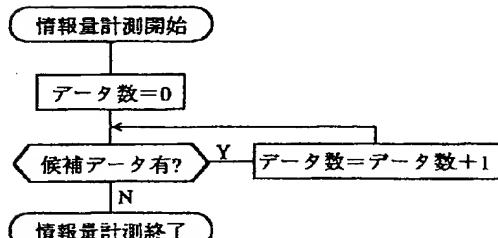
162 定型性情報保存部
164 定型性利用単語文字検索部

* 168 単語文字入力部
* 169 欄定型情報入力部

【図1】

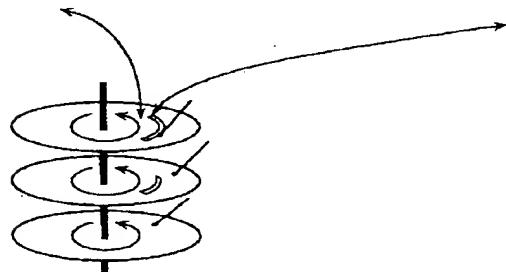


(2)

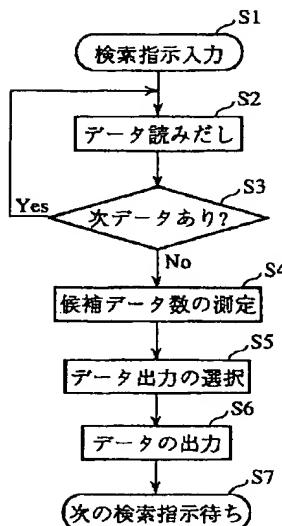


【図2】

タイトル							データ本体	
標題	入力時	属性	分類	著者等	キーワード	データ量	記憶場所	——
工業所有権法 逐条解説 (第12版)	95年 1月10日	コード (固有)	法律 工業 所有権法	編者 (特許庁) 発明協会刊	青本 特許 発明 ——	A5版 1200頁	HD001-0001 ~1200	----- -----
特許法概説 (第10版)	95年 1月11日	コード (固有)	——	——	——	——	——	——



【図4】



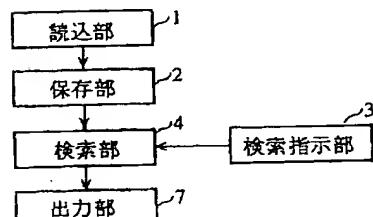
【図7】

データ内容	データ数
天気予報	5
会議議事録	2
日報	1
仕様書	1
テストサンプルデータ	30
計	39

【図10】

データ数	出力方法
10未満	データ内容表示
10以上30未満	データタイトル データ属性表示
30以上	データタイトル表示

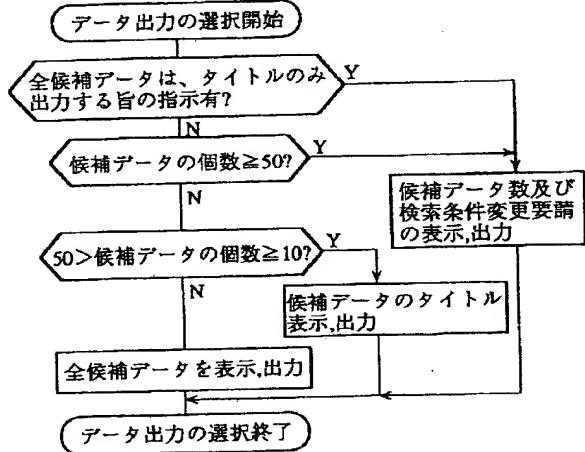
【図18】



【図3】

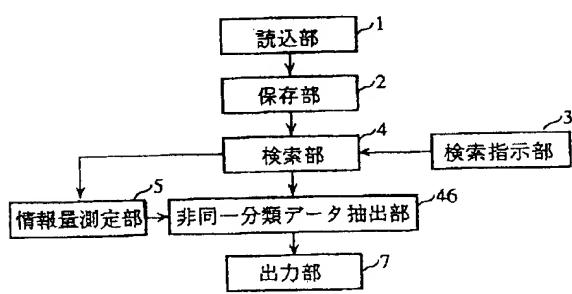
データ数	処理
10未満	データ内容表示
10以上50未満	データタイトル表示
50以上	検索条件を 変更して再検索

(1)

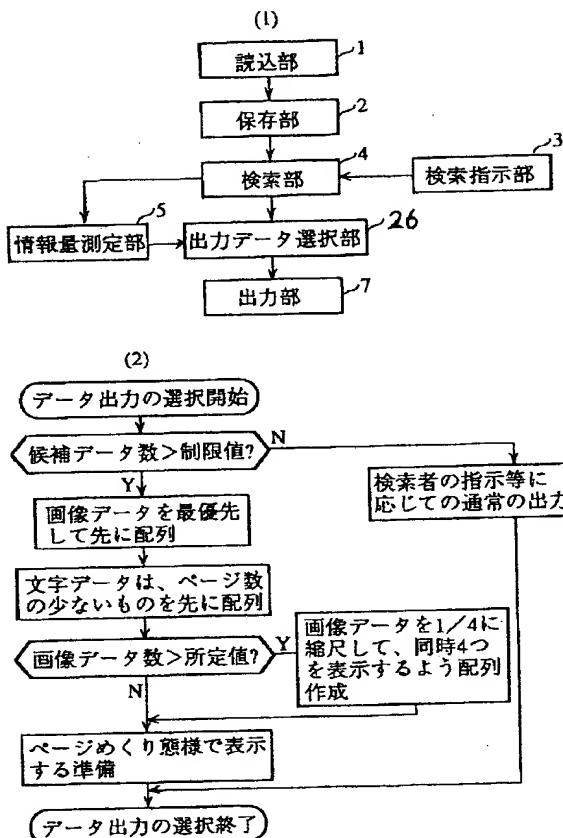


データ出力の選択終了

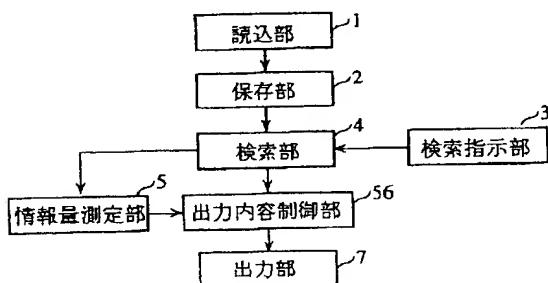
【図8】



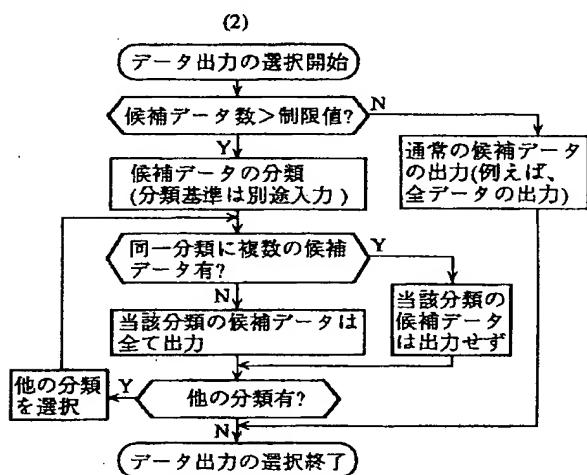
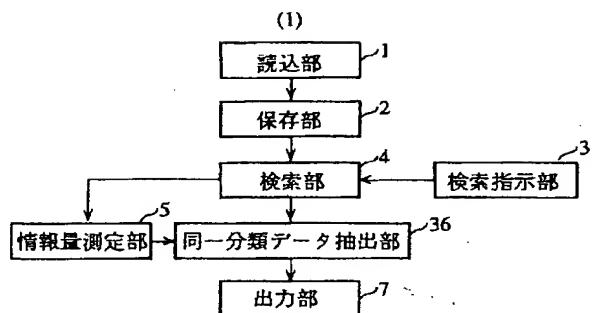
【図5】



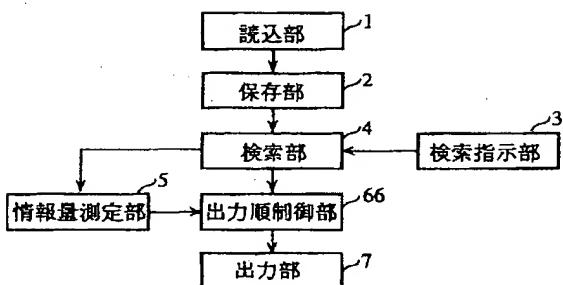
【図9】



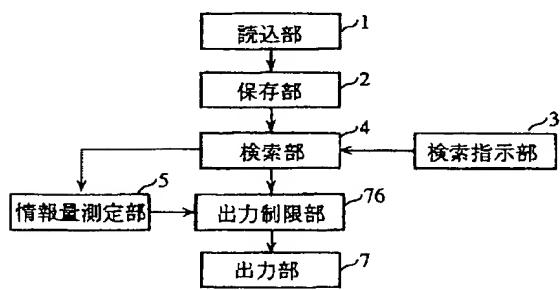
【図6】



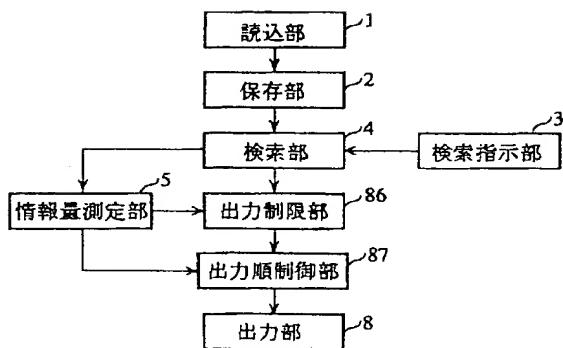
【図11】



【図12】

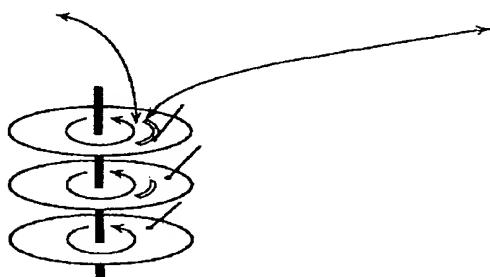


【図13】

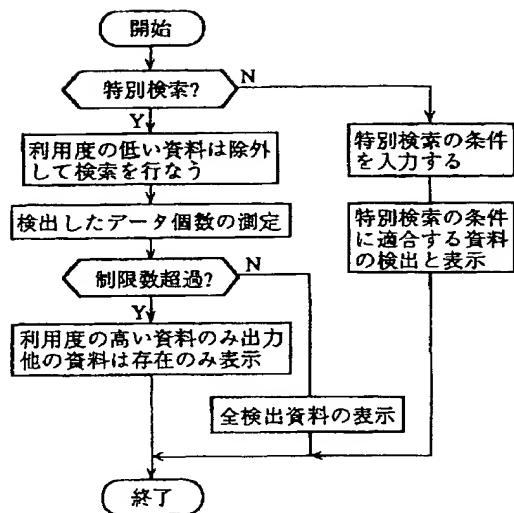


【図14】

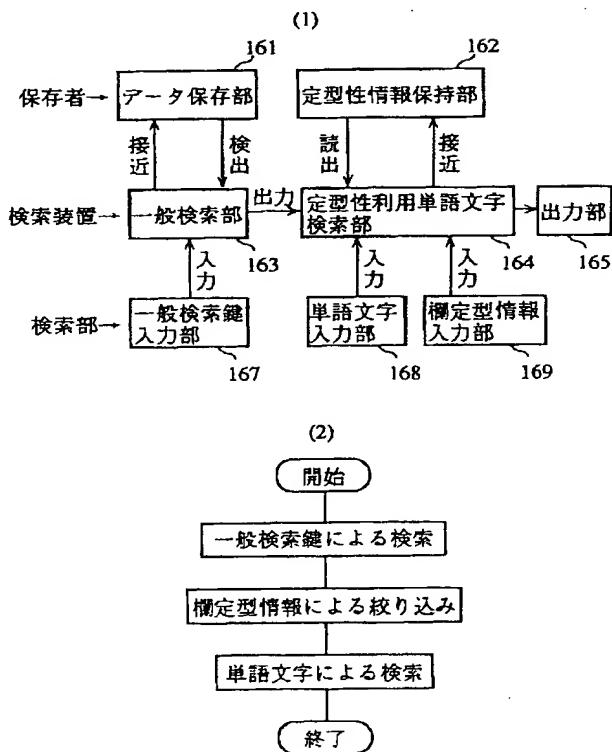
利用度 理由	タイトル								データ本体
	標題	入力時	属性	分類	著者等	キーワード	データ量	記憶場所	
大 (2)	工業所有権法 逐条解説 (第12版)	95年 1月10日	コード (国無)	法律 工業 所有権法	編者 (特許庁) 発明協会刊	青本 特許 発明	AS版 1200頁	HD001-0001 ~1200	----- -----
大 (1)	特許法概説 (第10版)	95年 1月11日	コード (国有)	-----	-----	-----	-----	-----	-----



【図15】



【図16】



【図17】

(1)
characteristic may be obtained.

What is claimed is:

1. A capacitor motor comprising a

(2)
Der Baum ist *so* hoch wie ich.

